AD-A239 219

ATION PAGE

Form Approved
OMB No. 0704-0188

(2)

| 2. DATE | 3. REPORT TYPE AND DATES COVERED |
|---|---|
| 07/23/91 | Annual technical report 3/1/90-2/28/91 |

**4. TITLE AND SUBTITLE**

Reading: Interactions with Memory

**5. FUNDING NUMBERS**

G-AFOSR 90-0246
PE 61102F
PR-2313
TA-A4

**6. AUTHOR(S)**

Gail McKoon

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**

Northwestern University
633 Clark
Evanston, IL 60208

**8. PERFORMING ORGANIZATION REPORT NUMBER**

AFOSR-TR- 91 0700

**9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)**

AFOSR/NL
Building 410
Bolling AFB D C 20332-6448

**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**

**11. SUPPLEMENTARY NOTES**

NA

ELECTE
AUG 0 8 1991
D

**12a. DISTRIBUTION/AVAILABILITY STATEMENT**

Approved for public release:
distribution unlimited.

**12b. DISTRIBUTION CODE**

**13. ABSTRACT** (Maximum 200 words)

In our research, we have proposed that the representation of linguistic and contextual information kept in working memory during reading or listening can be described as a discourse model, representing the concepts referenced in the discourse and the relations among them. The concepts in the model are hypothesized to vary in their accessibility, where accessibility is determined by the syntactic and pragmatic contexts in which the concepts are introduced. Subsequent reference to a concept is a function of the interaction of the expression used to reference the concept and the concept itself. The experiments in this report demonstrate support for the discourse model view by showing that difficulty of comprehension for pronouns and the degree to which they are fully understood depends on syntactic and pragmatic variables that affect the accessibility of the pronouns' intended referents.

91-07181

**14. SUBJECT TERMS**

Discourse comprehension; short-term memory; anaphora

**15. NUMBER OF PAGES**

56

**16. PRICE CODE**

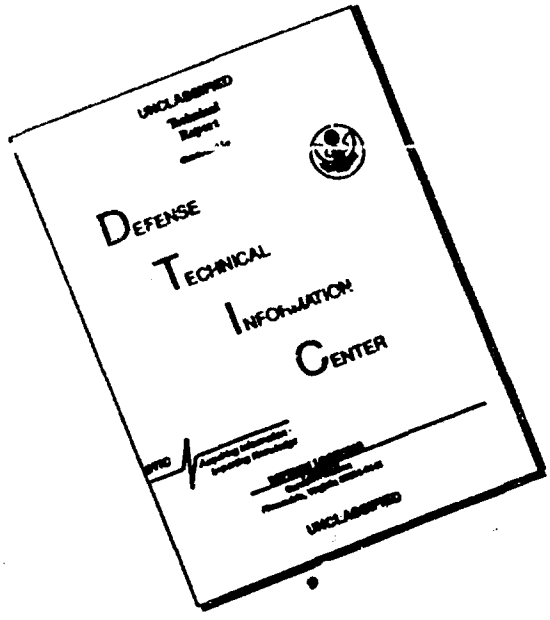
| 17. SECURITY CLASSIFICATION OF REPORT | 18. SECURITY CLASSIFICATION OF THIS PAGE | 19. SECURITY CLASSIFICATION OF ABSTRACT | 20. LIMITATION OF ABSTRACT |
|---|---|---|---|
| UNCLASSIFIED | UNCLASSIFIED | UNCLASSIFIED | Unlimited |



91

1991

# DISCLAIMER NOTICE




THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

# First Year Technical Report

Gail McKoon

Northwestern University

The main body of research in the first year was addressed to issues of discourse modeling. These issues concern the representation of a text that is built in short-term memory as the text is being read. Three papers have been written documenting this work, and these papers are summarized in this technical report. This work has been collaborative with Roger Ratcliff, Steven Greene, Gregory Ward, and Richard Sproat.

The second effort of the first year was to write a full description of the theoretical framework within which the current research is designed. This description has been accepted for publication in *Psychological Review*, and is enclosed with this report. This paper was written in collaboration with Roger Ratcliff.

## Discourse Models in Short-term Memory: I (McKoon, Ward, Ratcliff, Sproat)

When listening to a speaker or reading a text, information that has already been conveyed must be kept in mind as each new piece of information is understood. The context of prior information is used in determining the meanings of individual words, the relevance to the overall message of individual propositions, and the referents of pronouns and other referring expressions. In most current work in psycholinguistics, this information is assumed to be represented in 'working memory', and to have a rather simple structure. In our work, we propose that the structure is more complex, and that it is determined by a number of morphosyntactic and pragmatic factors. Following recent work in computational linguistics and discourse analysis, we label this structure a 'discourse model.' In several experiments, we investigate some of the referential properties of such a model. The experiments investigate the ease with which specific entities in the discourse model may be accessed by means of particular types of linguistic expressions, and show that successful reference is a function of both the pragmatic and syntactic context in which the referent was evoked in the prior discourse.

Within cognitive psychology, there have been two distinct traditions of text processing research that have investigated how on-line language comprehension relates to the representation of information in a discourse model. One tradition has generally focussed on syntactic determinants of linguistic structure, and, more narrowly, on structure within a single sentence. Within this view, the relationships among the elements of a sentence are organized according to the syntactic roles that they fill in that sentence. Reference to concepts or entities previously evoked by the text is accomplished by accessing syntactically defined elements; an anaphor accesses the syntactic part of the sentence in which its antecedent occurs. Ease of access is determined by the position of the antecedent in the syntactic structure. Mathews and Chodorow (1988), for example, provide data suggesting that antecedents more deeply embedded in a syntactic structure lead to more difficulty for the interpretation of an anaphor than antecedents not so deeply

embedded. In a similar vein, data from experiments by Nicol and Swinney (1989) suggest that the availability of a potential referent is a function of its 'syntactic appropriateness' as the antecedent of an anaphor. Syntactic approaches to the on-line representation of discourse information are reviewed by Mathews and Chodorow (1988) and by Fodor (1989).

The other traditional approach to on-line processing and representation has focussed on the structure of a discourse as a whole, rather than on single sentences. Under this approach, the basic units of analysis are (semantic) propositions, including the arguments of those propositions and the relations among those arguments. Kintsch (1974) originally proposed that a discourse was made up of 'individual idea units' and that these propositions were connected to each other through shared arguments. A connected set of propositions was assumed to consist of a 'topic proposition', i.e. the most important proposition of the set, and the importance of all other propositions was defined relative to this proposition. Kintsch and van Dijk (1978) later incorporated this structural proposal into a model of on-line comprehension. In this model, each new set of propositions in a discourse is added to the already existing structure via connections among shared arguments, with preference given to more recently mentioned propositions and arguments. Entities of the discourse that are more topical are more accessible because they are more likely to be kept active in short-term memory.

The 'discourse model' approach that we propose is similar to previous approaches in that it represents the entities evoked by a discourse and the relations among them. However, our model diverges from both of the traditional approaches in two crucial ways. First, we propose that the accessibility of discourse entities for subsequent anaphoric reference is determined not by syntax alone and not by topicality alone, but by a variety of syntactic, pragmatic, and semantic factors. The consequence of this proposal is that there need be no single, most accessible entity (such as the topic) in the discourse, nor is there a single metric (such as syntactic depth of embedding) by which accessibility can be calibrated. Second, we maintain that the accessibility of the entities in the discourse model is determined by an interaction between the context in which a particular entity was introduced into the discourse and the cue by which that entity is later accessed by the comprehension system. Different cues may access the same entity with varying degrees of success; in some contexts, a definite description may work better than a pronoun, and in other contexts, the reverse might be true. Furthermore, the entities that are most accessible given one cue may be different from the entities that are most accessible given another cue. For example, a pronoun may serve to evoke more recent entities, whereas a definite description might serve to evoke more distant entities.

Our theory of discourse models is based upon several key assumptions. First, following Sidner (1981), Webber (1979), and the propositional tradition (Kintsch, 1974), we assume that speakers (and writers) construct representations of the situations and objects that they wish to talk about. Such representations, or models, contain the entities ('arguments', Kintsch, 1974, or 'cognitive elements', Sidner, 1981) evoked in a discourse, and these entities are linked together by the relations in which they participate. One function of discourse, then, is to direct a hearer (or reader) to construct a conceptual model similar to that of the speaker (or writer). The degree to which the two models correspond will, in principle, determine the degree to which successful understanding is achieved. It is important to keep in mind that the entities in question are *conceptual* entities -- not linguistic ones. As Morgan (1978), Webber (1979), Sidner (1981), and others have pointed out, speakers use language, in particular referring expressions, to refer to objects in the world (or model thereof), and not to other linguistic units.

Second, we assume that the entire current discourse -- and not just individual component sentences -- is represented in the model. Certain subparts of the model are particularly salient, or 'in focus', at any given time (cf. Grosz, 1978; Grosz & Sidner, 1986). Although the whole discourse is represented in the model, at times, of course, portions of it will be relatively inaccessible. What constitutes the "whole" discourse is a matter of debate (cf McKoon & Ratcliff, submitted); many inferences of many sorts might be included, or only the most minimal. The critical point for the discussion of reference in this report is that the model contains all evoked entities, and the empirical questions concern the variables that control access to these entities.

Our third assumption is that the entities represented in the model are associated with varying degrees of accessibility. The accessibility of discourse entities is a function of a number of linguistic and non-linguistic factors including the morphosyntactic context in which an entity was introduced, and the relevance of the entity to the main topic of the discourse.

The varying degrees of accessibility of discourse entities, and the factors that govern accessibility, can be illustrated in a number of ways. For example, it has been claimed (Prince, 1981; Ward, 1985) that one of the functions of Topicalization (NP Preposing), illustrated in 1, is to mark the discourse entity represented by the preposed constituent as salient in the discourse:

(1) That painting, Mary paid more than twenty dollars for at the auction.

Given our view of discourse models, we would hypothesize that the preposed constituent in (1), *that*

*painting*, is used to evoke a discourse entity (corresponding to the painting for which Mary paid more than $20) that is more accessible than would have been the case had that NP *that painting* occurred in canonical post-verbal position.

Similarly, the effect of topicality on utterance interpretation can be illustrated with the two discourses in (2):

(2a). Sam likes the outdoor life. Having grown up in rural Kentucky, he knows a lot about nature and is an expert at fishing and shooting. He goes on hunting trips as often as he can. He used to hunt just small game, like rabbit and quail. However, lately he's taken up hunting deer. He thinks that they are really exciting to track.

(2b). Sam has many interests in the outdoors. He's an avid skier, and each winter he takes about a month off from work to ski in Colorado. In the summertime, he visits his parents in Montana where he has a chance to do some mountain climbing. Lately, for no apparent reason, he's taken up hunting deer. He thinks that they are really exciting to track.

In (2a), the topic of the discourse segment is hunting, whereas in (2b), the topic is outdoor sports. We would hypothesize that the discourse entity corresponding to the referent of the pronoun in each of the last sentences (i.e. *they/deer*) is more closely related to the topic of (2a) than that of (2b), and therefore we would hypothesize that it is more accessible in the former than the latter.

Our fourth assumption is that the accessibility of discourse entities changes as the discourse progresses, partly as a function of recency, and partly as a function of shifts in topic. Finally, our last assumption is that the accessibility of a discourse entity is determined by an interaction between the accessibility of that entity in a particular discourse and the linguistic cues used to refer to it. Consider, for example, the discourses in (3):

(3a). Mrs. Smith asked John to go to the grocery store to buy some groceries. However, John was tired, so he asked his roommate Mike to go instead. However, there was a slight problem -- she couldn't find the keys to the car.

(3b). Mrs. Smith asked John to go to the grocery store to buy some groceries. However, John was tired, so he asked his roommate Mike to go instead. However, there was a slight problem -- he couldn't find the keys to the car.

In both of these example discourses, Mrs. Smith and John are evoked by the initial sentence.

However, in (3a), the pronoun *she* seems to successfully access the discourse entity corresponding to Mrs. Smith, while in (3b), the pronoun *he* can be taken to specify either of the two more recently introduced male entities, Mike or John. Thus, even though Mrs. Smith and John are introduced at approximately the same point in the discourse, only the former is unambiguously accessible with a pronoun.

This theory of discourse models, with all the entities of the discourse represented with varying degrees of accessibility, and accessibility determined by a wide variety of factors, is a relatively new proposal in text processing research. Hence, there is little previous data that directly tests it. However, because the theory contains elements of previous approaches, it is consistent with previous empirical findings. In Kintsch's model for on-line text comprehension, the accessibility of an entity depends on the recency with which it was evoked and on how closely connected it is to the discourse topic. Empirically, both of these variables have been demonstrated to affect accessibility as hypothesized: it has been shown that more recently mentioned entities are more accessible (Jarvella, 1971; Caplan, 1972), and that entities more closely connected to the topic are bettered recalled (Kintsch & Keenan, 1973) and better recognized (McKoon, 1977). Because the discourse model theory incorporates both recency and topicality as variables affecting accessibility, these findings are consistent with it.

The theory of discourse models is also consistent with research motivated by more syntactic views of discourse representation. Under these views, the accessibility of an anaphor for an antecedent depends on the syntactic position of the antecedent. Matthews and Chodorow (1988), for example, tested comprehension of the pronoun in sentences like (4a) and (4b).

(4a). After the bartender served the patron, he got a big tip.

(4b). After the bartender served the patron, he left a big tip.

They found that reading times for the matrix sentences were faster when the antecedent of the pronoun occurred in subject position than when it occurred in object position. On a strictly syntactic account, this advantage would be due to a search process for the antecedent through the sentence's syntactic structure. An antecedent in subject position, as in (4a), would have an advantage in a left-to-right or top-down search. A discourse model approach would also predict an advantage when the antecedent is in subject position, but not because of a search through a syntactic structure. Instead, the advantage would be due to the greater accessibility of entities evoked in subject position relative to entities evoked in object position.

A number of studies have also investigated the use of syntactic structures in processing what have

been labeled 'syntactic gaps'. A gap is the 'empty' argument position in a sentence that results from the movement or deletion of that argument. In (5), the object of *accused* does not appear in direct object position; rather, it is related to the head of the relative clause (*skier*) via a long-distance syntactic dependency. Nicol and Swinney (1989) have presented evidence that the element which 'fills' the gap (e.g. *skier*) is quickly available at the gap site. They interpret this finding as a reflection of processes that depend on the syntactic structure of the sentence, i.e. the intended filler's syntactic status as the head of a relative clause. Another interpretation of the gap-filling data that is more in keeping with the discourse model approach is that the intended filler was made more accessible than other entities as the gap filler by virtue of its introduction as the head of a relative clause. It is this increased accessibility, and not syntactic status directly, that makes the filler available at the gap site.

(5). The journalists interviewed the skier that the waitress from the village accused of the crime.

In the discourse model approach, syntax is assumed to play a significant role in the interpretation of referring expressions both across and within sentences. Syntactic structure can function across sentences even though such structures would not typically be hypothesized to be available in working memory after a sentence boundary was crossed (Fodor, 1988; Mathews & Chodorow, 1988). In a discourse model, syntax does not affect referential processes directly; the search for a referent does not proceed through syntactic structures. Instead, it affects referential processes indirectly by determining the relative accessibility of entities in the discourse model. Several studies have investigated effects of syntactic structure outside of single sentences. Rothkopf, Biesenbach, and Billington (1986) and Rothkopf, Koether, and Billington (1988) have shown that a modifier is better recalled when it is presented in predicate adjective position than when it is presented in prenominal position. In Rothkopf et al.'s experiments, texts contained sentences with phrases like *the yellow fruit* or *the fruit that was yellow*. Subjects were better able to answer a later question about the color of the fruit if they had read the second (predicate adjective) version. Similarly, McKoon, Ward, and Ratcliff (in preparation) have shown that a predicate adjective is better recognized that a prenominal one. For example, the adjective *hostile* was presented in either prenominal or predicate position: *The hostile aunt was intolerant* or *The intolerant aunt was hostile*. Later recognition of the word *hostile* was faster and more accurate when it had been read in predicate adjective position.

While all of these experiments have provided results consistent with the discourse model theory, none was designed to test the theory directly. In the experiments described below, the notion of a discourse model was used to motivate empirical investigations of the accessibilities of the entities in a discourse.

Accessibility was examined through its effects on the ease of comprehension of pronouns; the more accessible an entity, the more easily comprehended should be a pronoun being used to refer to that entity.

Our goal was to test the hypothesis that accessibility depends on both the pragmatic and syntactic context in which an entity is introduced into a discourse. To vary the accessibility of an entity pragmatically, we manipulated how closely the entity was related to the overall topic of the discourse. To vary accessibility syntactically, we needed to find a pair of syntactically distinct yet semantically equivalent linguistic constructions so that we could vary accessibility while holding the meaning associated with the two constructions constant (or nearly so). That is, we needed two constructional variants that did not differ in meaning, one of which facilitated accessibility, the other of which inhibited it. A strong syntactic manipulation was necessary in order to examine the interaction of syntactic form with the pragmatic variable (topicality). According to the discourse model theory, the accessibility of an entity is not determined exclusively by the syntactic position in which it is introduced. Even if the entity is introduced in a syntactic position that is usually associated with a very low degree of accessibility, pragmatic factors would be able to greatly facilitate the accessibility of the entity.

The pair of syntactic constructions we chose were synthetic compound nouns and their non-compound (phrasal) equivalents. Synthetic compounds nouns in English, such as *deer hunter* and *French teacher*, consist of two parts: the right-hand member, or 'head' of the compound, typically a deverbal nominal, and a lefthand member that is interpretable as an argument of the deverbal nominal. The corresponding non-compound constructions, e.g. *hunts deer* and *teaches French* are verb phrase predicates that are semantically equivalent to the corresponding compound forms. In the compound variant, known in the linguistic literature as a type of 'anaphoric island', the compound-internal nominal (e.g. *deer* in *deer hunter)* has been argued to be categorically unavailable for subsequent anaphora (Postal, 1969; Simpson, 1983; Lieber, 1984). Our aim was to show that discourse entities evoked by such compound-internal nominals are, in fact, accessible in certain discourse contexts, and that their accessibility depends on discourse factors. In this way, the interaction of pragmatic and syntactic factors as predicted by discourse model theory would be demonstrated.

Anaphoric islands are particularly interesting from both a linguistic and psycholinguistic perspective, and in the next section of the introduction, we describe the island construction as well as its relevance to linguistic theory.

## Linguistic Background

It has often been argued that words are grammatically prohibited from containing antecedents for anaphoric elements, and thus constitute 'anaphoric islands' (Postal, 1969). In this context, *word* includes both those combinations of stem and affixes normally written as one word in English and also any compounds which may consist of several orthographic words in English. Specifically, Postal proposed the following principle of grammar on what he termed 'outbound anaphora': for any word W-1 no anaphor could have as an antecedent another word W-2 which is either "part of the sense of" W-1 or "morphologically related" to W-1. Thus, in a number of syntactic and morphological theories, it has been claimed that contrasts such as the one exhibited in (6) (Postal, 1969, p. 230) are the result of such a grammatical prohibition on outbound anaphora:

(6a). Hunters of animals tend to like them.

(6b). Animal hunters tend to like them.

Postal observed that the anaphor *them* may felicitously be used to refer to the animals in (6a), but not in (6b). The compound *animal hunters* in (6b) was claimed to render compound-internal elements (e.g. *animals*) unavailable for subsequent anaphoric reference. In this way, compounds participate in outbound anaphora by rendering compound-internal parts inaccessible for subsequent reference.

However, such analyses of anaphoric islands assume a particular -- and, we maintain, fundamentally incorrect -- view of reference. Specifically, anaphora is viewed as a relation between *linguistic* objects. For example, Postal's original formulation of the problem in terms of anaphoric islands involved morphosyntactic restrictions on possible antecedents for anaphoric elements: "Outbound anaphora is the relation between a *sentence* chunk, part of which is interpreted as antecedent, and some anaphor outside of that chunk" (Postal, 1969,p. 206). We claim that a more adequate account is possible once reference is viewed instead as a relation that holds between language and a discourse model (cf. Karttunen, 1976; Grosz, 1977; Webber, 1979; Sidner, 1979). Under this view, pronouns and other anaphors are used to refer to discourse entities, rather than to linguistic antecedents. The felicity of a particular instance of anaphora, then, is a function of the relative accessibility of the discourse entity to which the anaphor is used to refer.

In Ward, Sproat, and McKoon (1991), it is argued that outbound anaphora is not ruled out by any principle of grammar but is in fact fully grammatical and sensitive to the same types of pragmatic constraints as are other types of pronominal reference. Specifically, the felicity of anaphora involving nominal

compounds such as the one in (6b) is hypothesized to be a function of the accessibility of the discourse entity evoked by the word internal element to which the anaphor is used to refer. In those cases where the discourse entity evoked by the word internal antecedent is not sufficiently accessible, it is claimed that anaphora will be infelicitous.

In support of their analysis, Ward et al. present dozens of examples of naturally-occurring outbound anaphora, culled from a variety of oral and written sources. Consider the examples of outbound anaphora involving compound nouns in (7):

(7a.) *Bush* supporters would stay home, figuring *he*'d already won. (*he*=Bush) (J. Hirschberg in conversation, November 9, 1988)

(b). Call if you're a *small business* owner, or interested in starting *one*. (*one*==a small business) (TV ad, June 14, 1989)

(c.) *Museum* visitors can see through *its* big windows the 900-year-old Tower of London and the modern office blocks of the City financial district. (*its*=the museum's) (Associated Press Newswire, July 5, 1989)

In each case, the discourse entity evoked by the word internal antecedent is sufficiently accessible in these examples to permit subsequent pronominal anaphora.

However, outbound anaphora is not always felicitous, as evidenced by the deviance of (6b) above. Ward et al. (1991) claim that the difference in acceptability between (6a) and (6b) is not due to their structural difference per se, but rather to the pragmatic difference in accessibility associated with the relevant discourse entities. They suggest that the difference in accessibility may lie in the semantic difference between modifiers and predicates. First, they assume that compounds are instances of modifier-head constructions (see, for instance, Levi, 1978). That is, in the compound *Kal Kan cat*, *Kal Kan* can be said to modify *cat* in much the same way as the adjective *hostile* modifies *aunt* in the adjective-noun sequence *hostile aunt*. Furthermore, following Wilson and Sperber (1979), Rothkopf, Biesenbach, and Billington (1986), and Rothkopf, Koether, and Billington (1988), they assume that adjectives functioning as modifiers (in prenominal position, for example) are less salient than adjectives functioning as predicates. Given these assumptions, Ward et al. could account for the infelicity of many instances of outbound anaphora involving compounding with the following hypothesis: discourse entities evoked by modifiers are, ceteris paribus, less accessible than entities evoked by predicates and, therefore, that discourse entities

evoked in compound-internal syntactic positions are relatively less accessible than the same entities evoked in non-compound positions.

Thus, we have reason to believe both that the compound construction illustrated in (7) serves to render an entity relatively inaccessible in some discourse contexts but also that an entity evoked by a compound-internal element can be quite accessible in other discourse contexts. Specifically, when a discourse entity is topical, its accessibility is increased. This hypothesis regarding the relative contribution to accessibility of morphosyntactic and pragmatic factors makes a number of predictions amenable to empirical investigation.

In Experiment 1, we varied topicality for entities evoked by antecedents contained in both the compound and the corresponding verb phrase constructions, as shown in Table 1. Our prediction concerned the entity to which the pronoun in the final sentence of each text was intended to refer (*deer* in Table 1); we will refer to this entity as the "referent" entity. We predicted that the accessibility of this entity would be increased both by the pragmatic and the syntactic variables; the entity would be more accessible when it was more closely related to the topic and when it was introduced in a verb phrase rather than a compound.

## How to measure accessibility

In the discourse model theory, accessibility is defined as the ease with which a discourse entity, introduced at one point in a discourse, can be referenced at a later point in the discourse by some cue. The obvious cue with which to reference entities like the nominals considered in Table 1 is a pronoun. In all of the texts in Table 1, the pronoun of the final sentence (*they*) is intended as an anaphor for the target nominal of the preceding sentence (*deer*). The empirical goal is to measure accessibility by measuring ease of reference, that is, to measure the ease with which the pronouns in the final sentences are understood. This requires at least a minimal model of comprehension processes for pronouns.

In Greene, McKoon, and Ratcliff (manuscript) and Ward, Sproat, and McKoon (1991), we proposed that a pronoun is completely and correctly understood if its intended referent is sufficiently more highly accessible in the discourse model, relative to the pronoun as a cue, than all other discourse entities. Following current global memory models (Gillund & Shiffrin, 1984; Hintzman, 1988; Murdock, 1982; Ratcliff, 1978; see also Gernsbacher, 1989), a pronoun is assumed to be matched against all entities in the discourse model in parallel. The semantic and grammatical features of the pronoun are matched against the features of the discourse entities. Any particular entity in the discourse model will match the pronoun to

some degree, with the degree of match depending on both the entity's semantic and grammatical features as well as its accessibility. If the degree of match for some single entity is sufficiently high, and sufficiently higher than the match for all other entities, then that entity is identified as the pronoun's referent. If there is no entity that matches sufficiently well, then a referent is not identified. If more than one entity matches sufficiently (but none sufficiently better than the others), then again no single referent is identified. In the cases where a referent is not identified, comprehension may fail in the sense that the pronoun is left without a referent. Alternatively, selection of a referent might be postponed, waiting for more information from the discourse, or for strategic problem solving processes that might be able to identify a referent. In the usual case, where a single entity matches the pronoun sufficiently better than all other entities, the identification of the pronoun with the referent leads to the attachment in the discourse model of information associated with the pronoun to information associated with the referent.

This model for comprehension of pronouns makes the explicit claim that pronouns vary in the ease with which their referents can be identified such that, in some cases, no referent at all is automatically and uniquely identified. Failure to identify a unique referent might occur as the result of a number of factors, including the semantic and pragmatic content of the discourse and the speed required of comprehension processes by the speaker or reader. The possibility that pronouns sometimes fail to evoke unique referents has been discussed previously by Yule (1982), who points out that, in some discourse contexts, the identity of the entity referenced by an anaphor may be irrelevant to the reader or listener. Webber (1983) also suggests that, if there is no immediate need to determine a unique referent, an anaphor may be left unresolved. Empirically, failure to resolve pronouns has been demonstrated by Greene et al. (see below). Their experiments showed the difficulty of identifying a unique referent for a third person singular pronoun when two possible referents had been evoked in the discourse. Evidence for unique resolution was obtained only when reading rate was slow or readers could anticipate at exactly what point in the discourse the pronoun would occur. When reading rate was more normal (250 ms per word) or readers could not exactly anticipate the pronoun, the data suggested that no unique referent was identified.

The possibility that pronouns may sometimes be left unresolved complicates efforts to measure how difficult they are to comprehend. In particular, the time taken to read a pronoun (or the time to read a sentence containing a pronoun) is not an adequate measure. This is because reading times can reflect either time to successfully resolve a pronoun or time to process the pronoun but fail to resolve. One pronoun read in a given amount of time might be relatively easy to comprehend, and so be identified with a unique

referent, while another pronoun read in the same amount of time might be relatively difficult and left without a referent. In other words, reading time cannot be interpreted as a measure of comprehension difficulty unless it is combined with some method of determining whether the pronoun was successfully resolved. The method that has been typically adopted in previous research (cf. Chang, 1980; Corbett & Chang, 1983; Gernsbacher, 1989; McKoon & Ratcliff, 1980b) is to present the intended antecedent of the pronoun as a recognition test word at some point in the discourse after the pronoun. The reasoning that underlies this method is that successful resolution of the pronoun will increase the accessibility of its referent. This increase in accessibility will, in turn, facilitate the recognition decision about the referent when it is presented as a test word.

## Experiment 1

Table 1 shows the design of this experiment: the accessibility of a discourse entity was manipulated pragmatically, by how closely it was related to the topic of its text, and syntactically, by using either the verb phrase or the compound construction. The referent entity (*deer* in Table 1) was introduced in the next to last sentence of its text, and it was the intended referent of the pronoun mentioned in the last sentence. The hypothesis was that the accessibility of the entity would be increased when it was more closely related to the topic and when it was introduced in a verb phrase.

Subjects read each text one line at a time, in a self-paced procedure. After the final line, a single test word was presented for recognition (a decision as to whether or not the word had appeared in the text). For the texts that implemented the experimental design, the test word represented the intended referent of the pronoun in the final sentence. Increased accessibility for this referent was expected to result in faster reading time for the final sentence containing the pronoun, faster response time for the test word, or both.

Means for the reading times of the texts' final sentences and means for response times to the test words are presented in Table 2.

First, the data for the test words are considered. For each text, the test word was the referent noun, the antecedent of the pronoun in the final sentence. If, for all four conditions, subjects interpreted the pronoun correctly during the time they were reading the final sentence, then response times to the test word should be equal across the conditions. The processes of interpreting the pronoun might be more or less difficult across conditions, but if the correct referent was always evoked by the pronoun then it should be equally accessible across conditions at the time the test word was presented. This is what the data show:

there are no significant differences in response times to the test words (analyses of variance showed F's < 1.2 with both subjects and items as random variables).

Reading times show that there were differences in comprehension difficulty for the final sentences. It was hypothesized that interpretation of the pronoun would be difficult when the antecedent of the pronoun was in the modifier position in the compound. The data show this difficulty when the referent noun was low in topicality: reading times were longer when the noun was in a compound compared to when it was not. However, according to the discourse model theory, the difficulty should be reduced when the referent noun is more topical. This hypothesis was confirmed; topicality reduced reading times in the compound condition so that they were only slightly longer than in the verb phrase condition. These effects were supported by analyses of variance.

## Experiments 2 and 3

Our interpretation of the results of Experiment 1 depends on the assumption that subjects understood the correct referents of the pronouns in the final sentences of the texts in all of the experimental conditions. This assumption is consistent with the finding that response times for the test words were equal across experimental conditions. However, the assumption might be wrong. An alternative possibility is that the pronouns were not understood at all, and that this is the reason that response times to the test words did not differ across the experimental conditions. By this alternative, the differences in reading times would represent differing degrees of *unsuccessful* efforts at understanding the final sentences, and there would be no way to determine whether the same pattern of reading times would hold for successful efforts. Experiments 2 and 3 were designed to rule out this alternative.

In both of these experiments, the same texts were used as in Experiment 1. However, for each set of texts, two new final sentences were written. The new sentences were almost identical to the final sentences used in Experiment 1; one of the new sentences contained the same pronoun referring to the critical referent noun as in Experiment 1 and the other substituted for the pronoun a new noun that had not been mentioned previously in the text. For the set of texts shown in Table 1, the two new sentences were *And he says they are really exciting to track* and *And he says bears are really exciting to track*, where *bears* had not been mentioned previously in the text.

In Experiment 2, subjects read the texts with the referent noun in either its compound or its verb phrase form. The final sentence mentioned either the pronoun or the new noun. Following the final sentence,

the referent noun was presented as a test word. Response times for the test word can be predicted from the assumption that the pronouns in the final sentences are successfully resolved. If the pronoun in the final sentence is understood as referring to the referent noun, and it is this processing that leads to the facilitation of response times when the referent noun appears as a test word, then response times should be facilitated only when the final sentence contains the pronoun, and not when it mentions the new noun. This prediction was confirmed in Experiment 2.

Experiment 3 also used the two new versions of the final sentences, with the pronoun and with the new noun. However, two different test words were used: one was the referent noun and the other was a 'control' word picked from one of the earlier sentences of the text (e.g. *trips* for the texts in Table 1). Again, we predict response times to the test words from our assumption that the pronouns in the final sentences are understood to refer to the referent noun. This processing should lead to facilitation of response times for the referent noun test word but not for the control word, and the results confirmed this prediction.

## Experiments 4, 5, and 6

Experiments 1, 2, and 3, we argue, demonstrate that the time required to comprehend a pronoun is a function of the accessibility of the pronoun's referent in the discourse structure. When accessibility is reduced, either through the syntactic compound structure or through the context of the discourse, then comprehension takes longer. This was shown in the reading times of the sentences containing the pronouns. However, we pointed out that increased reading time does not by itself conclusively show that the pronouns were understood. In addition, some measure of the extent to which the pronouns were actually understood must be provided. Experiments 1, 2, and 3 used an immediate test of the antecedent of the pronoun (the referent noun) to provide evidence of comprehension. In Experiments 4, 5, and 6, we used a delayed priming procedure to provide the same kind of evidence.

The experiments involved a series of study-test lists. For each list, subjects read four texts, and then they were given a list of test words for recognition (responding positively if a test word had appeared in one of the studied texts, and negatively if it had not). For the experimental texts, the test words of interest were the referent noun (e.g. *deer*) and a modifier from the final sentence (e.g. *exciting*). These two words were presented in immediately adjacent positions in the test list, with *exciting* following *deer*, and so formed a 'priming' pair. From previous research (McKoon & Ratcliff, 1980a; Ratcliff & McKoon, 1978; Ratcliff & McKoon, 1988), it can be predicted that responses for the second word of the pair will be facilitated because they are closely related in memory by virtue of being from the same text.

In Experiment 4, the high topicality, compound versions of the texts were used. The final sentence of a text either contained a pronoun for which the referent noun was the correct antecedent (... *they were exciting to track*), or the final sentence contained the 'new noun' of Experiments 2 and 3 (... *bears were exciting to track*). With the pronoun version, the modifier *exciting* should be understood to apply to the referent noun; with the new noun version, it should not (it applies to *bears*). If subjects understand the final sentences completely, then *deer* will be more closely related in memory to *exciting* for the pronoun version of the final sentence than the new noun version. This increased relatedness should lead to greater facilitation of responses to *exciting* by *deer* when the final sentence contains the pronoun than when it contains the new noun. The results of the experiment followed this prediction as shown in Table 3.

In Experiments 5 and 6, the final sentence was always presented in the pronoun version. The questions were whether the referent noun and the modifier would be closely related in memory for both the high and low topicality versions of the text (Experiment 5) and for both the compound and verb phrase versions (Experiment 6). If the pronoun in the final sentence is understood to refer to the referent noun in all text versions, then the modifier and referent noun should be closely related for all versions. In other words, the amount of facilitation from the referent noun test word to the modifier test word should be the same in all cases; the results of the experiments confirmed these predictions.

## Discussion

A discourse model is the representation of information that is built during comprehension of a text or discourse. As comprehension proceeds through a text, the discourse model is continually updated and revised to include new input and to reflect the impact of new information on earlier information. The discourse model is made up of the entities evoked by linguistic and contextual information and the relations among the entities. The entities are assumed to vary in how salient they are in the discourse. Therefore, they vary in the degree to which they are accessible for later reference.

In this research, we propose that the accessibility of a discourse entity is a function of a number of factors, both linguistic and non-linguistic, arising from explicit information in the text as well as from contextual information, pragmatic knowledge, and speaker/writer and listener/reader goals. The accessibility of an entity for later reference is also determined by the cue with which it is referenced. A given entity may be quite accessible from one cue, but relatively inaccessible from another. Thus, accessibility is an interaction between entities in the discourse model and the cues used by the speaker/writer to evoke those entities.

The discourse model view of reference differs from previous approaches in three ways. First, we maintain that reference is a relation between the linguistic cue used to evoke an entity and the representation of that entity in the discourse model. That is, reference is not a relation between the linguistic cue and a linguistic antecedent, as has been assumed previously in both psychology and linguistics, by for example, Fodor (1989), Nicol and Swinney (1989), and Postal (1969). Second, we emphasize the interactiveness of the relation between linguistic cues and discourse entities. The accessibility of a discourse entity from a linguistic cue depends on the interaction of the accessibility of the entity in the discourse model in relation to the information (semantic, pragmatic, etc.) provided by the cue. And third, in our view, the accessibility of an entity for subsequent reference is a function not of its syntactic status alone (as in Fodor, 1989, Mathews & Chodorow, 1988, and Nicol & Swinney, 1989) and not of its pragmatic status alone (as in the topicality metric proposed by Kintsch, 1974). Instead, accessibility depends on both syntactic and pragmatic factors, and no doubt also on a variety of other contextual, semantic, and intonational factors.

The experiments presented here support the discourse model view by showing that both the morphosyntactic and the pragmatic context in which an entity is introduced into a discourse determine its accessibility for later reference. In Experiment 1, a referent entity (*deer*) was introduced in a morphosyntactic context that made it either more accessible (a verb phrase, *hunting deer*) or less accessible (a compound, *deer hunting*). Reading times for a sentence containing a pronominal anaphor for the referent entity were correspondingly faster when the entity had appeared in a verb phrase vs. a compound. The referent was also introduced in two pragmatic contexts; in one case, it was more closely related to the topic of its discourse than the other. Again, reading times for the sentence with the pronoun reflected accessibility, with faster reading times when the referent was more topical. In fact, the topicality effect was strong enough that reference in the compound condition was no more difficult than reference in the verb phrase condition: *they* as an anaphor for *deer* was no more difficult to understand when *deer* had been introduced in *deer hunting* than when it had been introduced in *hunting deer*.

These results validate the claim that the working memory used in text comprehension should be conceptualized as a discourse model. The results also support the claim that naturally-occurring productions of outbound anaphora, such as the examples given in the introduction to this report, are grammatically well-formed and that they are NEITHER 'performance errors' NOR the result of some type of pragmatic redemption (cf. Ward et al., 1991). The fact that examples of outbound anaphora are frequently produced in natural discourse does not necessarily entail that they are understood by the hearer/reader. But the

psycholinguistic data presented here indicates that they are, and that they are subject to the same types of pragmatic variables as are other kinds of anaphora. The pragmatic variable in our experiments, topicality, facilitated reference for both compound and non-compound constructions. Furthermore, placing the referent entity in a compound-internal position reduced its accessibility, just as a modifier position reduces the accessibility of other entities (McKoon, Ward, & Ratcliff, in preparation; Rothkopf et al., 1986; Rothkopf et al., 1988).

The results from these six experiments, taken as a whole, demonstrate the importance of using converging kinds of experimental data. It would not be possible for us to support our conclusions from measurements of sentence reading times alone. For example, we found that reading times were slowed when the referent entity for the pronoun in the final sentence was introduced within a compound. But we could have found that reading times in this condition were quite fast; this could have happened if the pronoun were uninterpretable and subjects quickly realized that it was uninterpretable. In this instance, the reading time data would have seemed to counter our hypotheses. However, an uninterpretable pronoun would be expected to lead to slow response times for the referent test word. Thus, only by simultaneous consideration of the sentence reading times and the test word response times can our interpretations of sentence reading times be justified.

As it happened, all the effects of degree of accessibility of the referent entity occurred in sentence reading times, and none in test word response times. That is, response times for the referent test word were not significantly different, whether the referent had been introduced with high or low topicality or in the verb phrase or compound structure. This raised the problem that subjects might not have understood the pronoun in any of the conditions. To address this problem, we used another version of the final sentence which made no reference to the referent entity being tested. If subjects did not understand the pronoun when it appeared in the final sentence, then it should make no difference to test word response times whether the sentence contained the pronoun or not. But, it did make a difference; response times for the referent test word were faster when the final sentence contained the coreferential pronoun than when it did not. It appears that subjects did understand the pronouns in the final sentences, and that, consequently, reading times for the final sentences did reflect comprehension difficulty.

Additional confirmation that the pronouns were understood as specifying their intended referents was provided by the last three experiments reported here. In these experiments, testing was delayed; testing for a text was separated from reading the text by other texts and test items. With this procedure, testing

examines the representation left in memory after reading and comprehension processes are complete, unlike immediate testing which examines memory in the midst of comprehension. With immediate testing, results might reflect transient processes and not information that is encoded into the long-term memory representation of the text. These transient processes might seem to indicate that the pronoun was understood even when it was not understood completely enough to effect connections among the appropriate pieces of information. Complete comprehension of the pronoun should lead to connections among pieces of information that were attributed to the referent and pieces of information that were attributed to the pronoun. Experiments 4, 5, and 6 confirmed that these connections were encoded. The experiments showed that a modifier (*exciting*) attributed to the pronoun (*they*) was closely related to the intended referent (*deer*). The convergence of this result with the immediate testing results from Experiments 1, 2, and 3 demonstrates once again that readers did comprehend the pronouns as having their intended referent.

Through these converging sets of data, we argue that the difficulty of comprehension for a pronoun depends on the accessibility of the discourse entity to which the pronoun is being used to refer. Pronoun comprehension is not viewed as a process that depends on the pronoun alone or even primarily. The issue for the comprehension system is not how to use a pronoun to access the intended referent. Instead, the issue is how the discourse model is constructed from the discourse in such a way that pronouns are automatically and correctly interpreted.

## Discourse Models in Short-term Memory: II (McKoon, Greene, Ratcliff)

When we encounter a pronoun in a discourse, we usually feel as if we understand its referent immediately (cf. Clark & Sengul, 1979). We are not consciously aware of any pronoun resolution mechanism operating, or of any disambiguation strategies that we might employ. Because of this unawareness, most psycholinguists studying pronominal reference have been tempted to assume that the psychological process involved is automatic. That is, researchers have implicitly assumed that the process under investigation in studies of pronoun resolution is always triggered when a reader encounters a pronoun and that the process is always carried through to completion--the identification of a unique referent for every pronoun. The questions for recent research have been how soon after the occurrence of the pronoun is the process triggered and how many possible referents are considered (cf. Chang, 1980; Corbett & Chang, 1983; Gernsbacher, 1989). Unfortunately, fifteen years of research based on the belief that pronominal referents are automatically identified have so far failed to produce a satisfactory account of the process of pronoun resolution.

In this research, we propose a new framework within which to view the process of pronoun resolution. This framework is motivated by both empirical and theoretical considerations. First, we take seriously the notion of an automatic process (Neely, 1977; Posner & Snyder, 1975; Ratcliff & McKoon, 1981). Previous research on pronoun resolution has left the assumption of automaticity implicit and, thus, untested. One goal of the present research is to state explicitly what is automatic and what is strategic in pronoun resolution, and to subject these claims to empirical verification. More importantly, our theoretical framework draws from contemporary work in discourse representation and in global memory models, as outlined in section I above. Whereas early theories of discourse comprehension were based on the verbal learning tradition and modeled discourse as a single dimensioned list of clauses or propositions, ordered serially or hierarchically (e.g., Clark & Sengul, 1979; Jarvella, 1971; Kintsch, 1974), recent discourse models organize information in multidimensional ways that more strongly reflect local context (e.g., Grosz, Joshi, & Weinstein, 1983; Webber, 1983). Similarly, most of the early process models for identifying referents of pronouns employed, either explicitly or implicitly, a serial linear or hierarchical search (e.g., Clark & Sengul, 1979; Corbett & Chang, 1983; Hobbs, 1978; van Dijk & Kintsch, 1983; see Matthews & Chodorow, 1988, for a review). These models were inspired by the memory scanning retrieval models of the time (e.g., serial scanning models, Murdock, 1974), which have now largely been replaced by global parallel retrieval models (e.g., Gillund & Shiffrin, 1984; Hintzman, 1988; Murdock, 1982; Ratcliff, 1978).

Hence, we replace the metaphor of the pronoun as a trigger initiating a serial search through a minimally structured textual representation with that of the pronoun as a cue to the most likely entity in a rich discourse representation.

Viewed in this way, the problem for research is not to investigate the mechanics of how a search process triggered by a pronoun might proceed, but instead to investigate how a discourse model is constructed during comprehension so as to make the use of pronouns felicitous. Each entity in a discourse is assumed to have some degree of accessibility, which is determined in part by the syntactic and semantic structures in which it is linguistically expressed. Accessibility is measured relative to the local environment, that is, relative to the other entities introduced in nearby clauses and sentences. As the reader or listener moves through a discourse, the accessibility of entities changes as the local environment changes. The entity or entities that are most accessible at any point are what the discourse is about at that point--a notion that various authors have attempted to capture in the concepts of a discourse segment's "focus" (Sidner, 1983b), "center(s)" (Grosz et al., 1983), or "topic" (Reinhart, 1982), and which we will refer to by the term "focus of attention."

Semantic and pragmatic factors also contribute to the relative accessibilities of discourse entities. For example, the perceived causal agent of a verb may be more accessible than its other arguments (Hudson, Tanenhaus, & Dell, 1987), and a discourse entity may be more accessible if it is more closely related to the topic of its discourse (McKoon et al., 1990). In addition, changes in relative accessibility can be signaled by certain conventional words and phrases that are used to indicate a shift in discourse focus (Grosz, 1981).

The accessibility of entities in a discourse is determined not only by the local environment at the time they are initially introduced, but also by subsequent reference to them or to objects or properties associated with them. For example, noun anaphors can increase the accessibility not only of the concept to which they refer, but also of other concepts that were mentioned in the same clause as the noun with which they corefer (Dell, McKoon, & Ratcliff, 1983). Certain concepts also permit the use of "associative anaphora" (Hawkins, 1977): after introducing the topic of a car, a reference to "*the* steering wheel" is felicitous. The initial reference to the car makes its parts accessible enough that they can be referred to using the definite article, usually reserved for previously mentioned entities (see also Chafe, 1976; Clark & Marshall, 1981; Prince, 1981).

The framework we put forward here is intended to suggest how referents for pronouns can be identified in the context of a highly structured discourse model, rather than the simple linear representation

implicit in previous research (e.g., Clark & Sengul, 1979; Corbett & Chang, 1983). In our framework, a pronoun must be evaluated against the rich and complex structure established by the syntactic, semantic, and pragmatic factors that determine the relative accessibilities of the different entities in the discourse. We propose that a pronoun can be completely and correctly understood if its intended referent is sufficiently more highly accessible in the comprehender's discourse model, relative to the pronoun as a cue, than all other discourse entities. We base the process by which a pronoun is matched against possible referents on current global memory models, as described in section I (Gillund & Shiffrin, 1984; Hintzman, 1988; Murdock, 1982; Ratcliff, 1978; see also Gernsbacher, 1989). In the proposed process, the semantic and grammatical features provided by an anaphor (as a retrieval cue) are matched automatically and in parallel against the semantic features of all entities in the current discourse model. A particular entity will match the anaphor to some degree depending on how accessible the entity is from the anaphor as a cue. Both the features of the entity (e.g., gender and number) and its accessibility will contribute to a determination of the degree to which it matches. If the degree of match for a single discourse entity is sufficiently high and better than the match for all others entities, that entity is automatically identified as the anaphor's referent. If there is no entity that matches sufficiently well, then no referent is identified, and selection of a referent is postponed or some kind of strategic (problem solving) process can be invoked. If more than one entity matches sufficiently, then again selection is postponed to wait for more content from the discourse, or strategic problem solving can be attempted. In the usual case, when one entity matches sufficiently better than all others, the information in the propositions that include the anaphor is combined with the information from the propositions that include the referent entity.

Hence, in this framework, pronouns are resolved either by an automatic matching process or, if that process fails to produce a discourse entity that matches the pronoun sufficiently better than all other entities, an optional strategic process. This account of the mechanism by which pronouns cue potential referents can be applied to a variety of different discourse contexts. Most often, a pronoun is used to refer to a single discourse entity that is already easily accessible based on the syntactic and semantic context in which it was introduced-- an entity that is in the reader's or listener's focus of attention (Brennan, 1989; Chafe, 1974; Fletcher, 1984; see also Givon, 1976). In this situation, the pronoun matches a focused entity to a high degree and sufficiently better than all other entities in memory. As a result, the propositions that include the pronoun can be simply and automatically attached to the entity that is in focus at the time of the pronoun's use, with the consequence that the accessibility of the focused entity is maintained or enhanced. Pronouns are usually used when the focus of attention of the discourse has not shifted (Grosz et al., 1983), so the

default procedure of attaching new propositions to focused entities may have little processing cost.

Although pronouns may often be used to refer to a single, most accessible, entity, a processing model in which a pronoun can vary in the degree to which it matches previously evoked entities leads directly to the possibility that sometimes there may be no discourse entity that matches sufficiently better than all others. This could come about either because no entity matches well, or because several entities match about equally well. In these cases, no referent is automatically and uniquely identified for the pronoun. Various factors, such as the reader's or speaker's speed, the reader's or listener's comprehension goals, and the surrounding discourse context may conspire to make this possibility more or less likely. Variations in these factors can affect the degree to which a pronoun evokes its intended referent, so that in some contextual conditions a pronoun will succeed in matching its intended referent, while in others it may fail to do so. In the case in which no referent matches sufficiently well and strategic processes are not invoked, then no referent will be identified and there may be no effect on the relative accessibilities of discourse entities as a result of reading the pronoun. In the case in which several entities all match well, and about equally well, then they may collectively represent the focus of attention. None of them would be singled out as a unique referent, and information about the pronoun would be attached to them jointly as the focus, with the consequence that their relative accessibilities would not change as a result of reading the pronoun.

The possibility that people might sometimes fail to identify unique referents for pronouns has been suggested in the linguistic literature. Emphasizing the need to take the comprehender's purposes into account, Yule (1982) argues that comprehenders will sometimes interpret the discourse "in terms of some information marked for attention predicated of some individual or group, the referential identity of which is not an issue" (p. 319). Webber (1983) makes a similar point-- if there is no single best matching discourse entity for an anaphor, and if there is no immediate need to choose a referent for the anaphor, then the comprehender may simply leave the reference unresolved. If readers or listeners have little inducement to identify the referent of a pronoun, they may simply associate the information from the propositions that include the pronoun with whatever entities are currently accessible.

Our proposal, that anaphoric processing involves an automatic matching process that may sometimes fail to produce a referent, cannot be evaluated with respect to past research in any simple way. In the earliest studies of anaphoric reference (cf. Clark & Sengul, 1979; Haviland & Clark, 1974), it was assumed that the referents of pronouns would always be identified (probably a correct assumption for the

texts that were used), and the exact point at which identification took place was not at issue. The only question was how difficult the identification process would be, and difficulty was measured by reading time. The more difficult the identification process for a pronoun in a sentence, the longer would be the reading time for the sentence. In more recent studies, the questions at issue have changed to focus on whether, and when in the time course of processing, a referent for an anaphor is understood (Chang, 1980; Corbett & Chang, 1983; Dell, McKoon, & Ratcliff, 1983; Ehrlich & Rayner, 1983; McKoon & Ratcliff, 1980, 1981, 1984; Nicol & Swinney, 1989; Tanenhaus, Carlson, & Trueswell, 1989). The results of these studies still do not lead to a direct test of our proposal, but the studies do offer an appropriate methodology. We first explain the methodology, and then consider the possible implications of previous results.

The procedure introduced by Chang (1980; also Caplan, 1972) was a probe task, in which possible referents of a pronoun are presented as test words for recognition. Subjects read or listen to a short discourse that describes two characters and then refers unambiguously to one of them with a pronoun. At some point after the pronoun, the subject is shown a character's name and asked to verify that the character was mentioned in the discourse just presented. The tested name can be either the intended referent, the other character, or some name that was not in the discourse at all. For example, in the final sentence in Table 4, the pronoun *she* is intended to refer to Mary, and either Mary, John, or some other name could be presented as a test word. For the character names that are in the discourse, the correct response is *yes*, the name was mentioned in the discourse. The result that was always expected by previous researchers is that responses to the name of the intended referent, *Mary* in Table 4, will be faster and more accurate than responses to the name of the other character, *John*. The reasoning is that the processes by which the pronoun is understood leave the intended referent in a more accessible state than the other possible referent, and this increased accessibility leads to relative facilitation for the referent as a test word.

Our proposed framework differs from previous views in the claim that the unique referent of a pronoun may or may not be identified, depending on contextual conditions. Under some conditions, the automatic process of matching the features of a pronoun against the features of entities in memory will succeed in producing a discourse entity that matches the pronoun sufficiently better than other entities, and so the referent of a pronoun will be uniquely identified. The result will be to leave the identified referent in a state of high accessibility that will, in turn, lead to relative facilitation when the referent is presented as a test word. But under other conditions, the process may fail to identify uniquely the intended referent, and then its accessibility will not be high relative to the accessibilities of other possible referents, with no

resulting facilitation for the intended referent relative to other test words.

Tests of this proposed framework depend critically on the assumption that the matching process of pronoun resolution is relatively fast and automatic. This assumption is adopted because it accords with our intuition that pronouns are normally processed quickly and effortlessly. We make this assumption explicitly in order to distinguish the automatic matching process from other, more strategic and usually slower processes that might come into play if a single, best matching entity is not produced.

In many previous studies that have used the probe word procedure to investigate pronoun comprehension, reading times and response times have been slow enough that it is doubtful whether automatic processing could be claimed. Since Chang (1980) first used the test word procedure to investigate pronoun comprehension, others have followed (Corbett & Chang, 1983; Gernsbacher, 1989), with a virtually unanimous result: responses to the intended referent presented as a test word are facilitated relative to responses for other possible referents presented as test words. But in each case, either reading times or response times or both seem slow. For example, Corbett and Chang (1983; Experiment 1) found faster responses for the intended referent than another possible referent, but response times were slow (800-900 ms) and so were reading times (about 380 ms per word, controlled by the subjects). Gernsbacher (1989) used reading times of over 500 ms per word (controlled by the experimenter), with response times in the 1000 ms range. In addition, previous studies may have encouraged strategic processing of pronouns, not only by using slow reading rates but also by a specific task demand-- asking for the identity of the pronoun immediately after reading. For example, for the text in Table 4, subjects would be asked "who dropped it on the counter?" immediately after reading the text. The motivation provided by such a specific question in combination with a reading rate slow enough to give time to answer the question during reading may have led subjects to adopt strategies that they might not have under other task conditions.

Our goal for the experiments described here was to examine pronoun comprehension as an automatic process. To accomplish this, we changed the experimental procedures used in previous research in two ways. First, both the reading rate and the time for responding to the test word were speeded relative to previous experiments. Second, we eliminated task demands that might encourage special strategic processing of pronouns, such as questions about the identity of a pronoun. Both of these changes were motivated from general notions about automatic processing developed in research areas other than reading (cf Posner & Snyder, 1975; Ratcliff & McKoon, 1981), and the application of these notions to reading is not straightforward. However, as will be seen, the procedural changes brought about substantial changes in

experimental results, lending support to the application of an automatic/strategic distinction to investigations of reading processes.

The procedural changes designed to speed reading and response times were guided both by findings from other research domains and by intuition. What times qualify as within the range of automatic processes is fairly clear for recognition responses, from both Posner and Snyder's (1975) original studies and a number of other studies with various methodologies (e.g., Neely, 1977; Ratcliff & McKoon, 1981). However, for reading time, deciding what rates qualify as automatic presents a problem; it is not clear how automatic reading processes can be separated empirically from slower, strategic reading processes, or even whether there is a clearly separable dichotomy between the two kinds of processing in reading. What we decided to do was to speed up the presentation rate of our materials from the rates used by earlier researchers to a rate more nearly approaching what college students have been estimated to use normally. Using texts considerably more difficult than those in the experiments presented here, other researchers (e.g., Just & Carpenter, 1980; Rayner, 1978) have found average reading speeds in the range of 200 to 250 ms/word. For texts more similar to those in the following experiments, Ehrlich (1983) found mean eye fixation times consistently below 300 ms, but because only about two thirds of the words of a typical text are actually fixated (Just & Carpenter, 1987), one can calculate the effective reading speed to be about 200 ms/word. In fact, Just and Carpenter (1987) consider a reading rate of 240 words per min, or 250 ms/word, to be "normal" (p. 38). So, in our experiments, we set the reading rate at 250 ms per word. We also instructed subjects to respond quickly with high accuracy, with the intention that response times should be in the 700 ms range. Based on past experiments (e.g., Dell et al., 1983; McKoon & Ratcliff, 1986, 1989b), we expected that subjects would be able to achieve this level of performance.

The materials in our experiments were modeled on those typical of previous studies of pronouns (Chang, 1980; Corbett & Chang, 1983; Gernsbacher, 1989), except that we used longer texts. Each text began with a sentence that introduced two characters with proper names, continued with a sentence that did not emphasize either character, and concluded with a final sentence made up of two clauses. In the first of these clauses, both characters' names were mentioned (in the same order as in the first sentence), and in the second clause, there was a pronoun intended to refer to the first-mentioned character (the subject of the first clause). The pronominal reference was unambiguous, both because the sex of the two characters differed and because the predicate of the second clause described an action that could only be performed by the referent character. An example of one of the texts is shown in Table 4.

In a discourse model of this text, the two characters would be of about the same accessibility. Both were introduced at the beginning of the discourse and both were rementioned in the first clause of the final sentence. However, the first-mentioned character might enjoy a slight advantage simply because of being mentioned first (Gernsbacher & Hargreaves, 1988; Gernsbacher, Hargreaves, & Beeman, 1989). Also, the first-mentioned character was the subject of the first clause of the final sentence, and the grammatical subject of a sentence is a good candidate for coreference with a subsequent subject-position pronoun (Matthews & Chodorow, 1988; Sidner, 1983b). So, before the reader encounters the pronoun, the first-mentioned, subject character may be more accessible than the object character. This initially higher accessibility might lead to a sufficiently higher match between the subject character and the pronoun, assuming also a match in gender and number, that the subject character is identified as the referent. As a result, the propositions that include the pronoun would be attached to those that include the subject character. The processing involved in attaching the propositions might further increase the referent's accessibility, giving an advantage to the referent when it is presented as a test word.

On the other hand, the grammatical subject might not have an advantage over the grammatical object. The object of a verb in the main clause of a sentence is also often a good candidate for subsequent pronominalization (Clifton & Ferreira, 1987; Sidner, 1983a). Thus, the subject and object might not differ in accessibility; they might both be in the reader's focus of attention. If, in an experimental context, gender does not sufficiently distinguish the two characters, the match between the pronoun and the intended referent might not be sufficiently higher than that between the pronoun and the non-referent. In this situation, subjects could engage in further, possibly strategic, processing to choose between the possible referents. Or they could simply attach the new propositions to the discourse entities currently in the focus of attention, not distinguishing between the two possible referents because they are both in the focus of attention. In this case, processing of the pronoun would give no advantage in accessibility to either of the two characters over the other.

Experiments 1 and 2 were designed to distinguish between the two hypotheses just described: the subject character might have an advantage in accessibility such that it is identified as the referent of the pronoun and therefore given an increase in accessibility, or it might be that neither character has a sufficiently great advantage to be uniquely identified as the referent, and so neither would gain in relative accessibility. The first hypothesis predicts that processing of the pronoun will facilitate responses to the intended referent relative to responses to the other character name, whereas the second hypothesis predicts

that there will be no facilitation of the referent relative to the other character. If the second hypothesis is upheld, it would suggest that readers do not always identify a unique referent each time they encounter a pronoun.

The experiments below suggest that readers do not, in fact, always automatically identify referents for pronouns. In Experiments 1 and 2, processing of the pronoun did not facilitate responses to the referent test word relative to the nonreferent test word. Because this is a null result, we conducted a further seven experiments. Experiments 3 and 4 added more subjects and used pronouns for which the intended referent was the object instead of the subject of the first clause of the final sentence. There was still no relative advantage of referent test words over nonreferent test words. Experiments 5, 6, and 7 compared our procedure (relatively fast reading times and relatively fast responses) to a procedure with much slower reading times and responses times that has previously been shown to produce facilitation of referents relative to nonreferents (Gernsbacher, 1989). With the slow procedure, we did find facilitation of referents relative to nonreferents, but only when the experimental texts were short enough that subjects could predict the occurrence of the pronoun and the test word. This pattern suggests that our finding of no relative facilitation of referents differed from past findings of facilitation because of the difference in procedures and materials. We argue that, with the slow procedure and the predictable materials, subjects invoke strategic processes to resolve the pronoun references. Finally, in Experiments 8 and 9, we used the fast procedure to compare comprehension of the pronouns to comprehension of nominal anaphors. We replicated what has previously been shown (Dell et al., 1983), that processing of a nominal anaphor, such as *the criminal*, facilitates responses for its referent (*burglar*) and responses for words associated in the text with the referent. Thus, we show that our fast presentation rate is not so fast that it prevents all types of anaphoric processing. In the discussion section, we argue that automatic processing of anaphors does occur with our fast procedure, as evidenced by the results for nominal anaphors, but that automatic processing does not identify a single best referent for the pronouns under investigation. Instead, the propositions that include the pronoun are simply attached to the entities in the focus of attention at that point in the discourse. Because the texts used in these experiments leave both the referent and the nonreferent characters in the focus of attention, neither is given an advantage over the other.

## Experiments 1 and 2

An example of the texts used in these experiments appears in Table 4. As described above, the first sentence introduced two characters of different gender, the second sentence did not emphasize either

character, and the final sentence consisted of two clauses. The first clause of the final sentence had one of the characters as subject, the other as object, and the second clause referred to the subject character with a pronoun. The words of the texts were presented on a CRT screen, one at a time at the rate of 250 ms per word. When a test word was presented for recognition, all preceding words of the text were erased from the screen, and subjects were instructed to respond *yes* if the test word had appeared in the text just presented, and *no* if it had not.

The aim of the experiments was to determine whether processing of the pronoun gave a relative advantage in accessibility to the referent character. Exactly how to design experiments to address issues like this has been the subject of considerable discussion (cf. Dell et al., 1983; MacDonald & MacWhinney, 1990). It is, first of all, important to distinguish two different questions that might be asked: one is whether the referent has an advantage relative to the nonreferent, and the other is whether the referent has an advantage relative to some neutral baseline. We were mainly concerned with the first question, for which the choice of experimental design is straightforward. To find out whether processing of the pronoun gives a relative advantage to the referent test word, we compared responses to the referent and nonreferent test words when the test words were presented before the pronoun to responses when the test words were presented after the pronoun. If processing of the pronoun gives an advantage to the referent, then whatever difference there was in referent and nonreferent responses before the pronoun ought to change in the direction of relative facilitation for the referent. There might, of course, be changes in baseline response time or accuracy as the test point is changed from before the pronoun to after the pronoun, but this would be a simple main effect that should not obscure any change in the relative differences of referent versus nonreferent responses.

We implemented this design in Experiment 1 with two test positions for the referent and nonreferent test words. One test position was immediately before the pronoun in the final clause, and the other was after the word following the pronoun; these are test positions 1 and 2 in Table 4. With the text presented at 250 ms per word, the test at position 2 occurred 500 ms after the pronoun was displayed. Experiment 2 was the same as Experiment 1 except that the two test positions were immediately before the pronoun and at the end of the final clause, test positions 1 and 3.

Although we were mainly interested in the relative facilitation given by processing of the pronoun to the referent and nonreferent characters, we also included in the design a test of a hypothesis put forward by Gernsbacher (1989). She proposed that processing of a pronoun gives relative facilitation to the referent

test word by means of suppressing the accessibility of nonreferents. As support for this hypothesis, she showed that response times to a nonreferent test word slowed at the end of a sentence containing a pronoun while response times for the referent test word stayed about the same as before the pronoun (Gernsbacher, 1989; Experiment 3). To test her hypothesis, we included a "control" test word in Experiments 1 and 2. This was a word that had appeared in the text in the first or second sentence (so the correct response for recognition was *yes*, the same as for the referent and nonreferent test words). By presenting this word at the same two test points as the referent and nonreferent test words, we could trace changes in response times that should be independent of effects of processing the pronoun. For example, it might be that responses for all test words are slower at the end of a sentence than in the middle of a sentence because the end-of-sentence test word is competing for processing capacity with end-of-sentence comprehension processing. If this were the case, then further research would be needed to support the suppression hypothesis.

It is important to note that the control test word was included only to address the suppression hypothesis. Neither the control word nor any combination of the conditions in the experiment allows the issue of true facilitation relative to a neutral baseline to be addressed. As was pointed out, this issue is not directly relevant to the hypotheses of concern here.

Means for each condition are shown in Table 5. In Experiment 1, with test positions 1 and 2, there were no significant differences between the test positions. The only significant effect was for test word, with slow responses for the control test word. The pattern of results was similar for Experiment 2, in that there were no significant differences between the referent and nonreferent test words as a function of test position. The effect of test word was significant, as was the effect of test position, and the interaction of the two variables. The significant interaction is due to the difference between the control test word and the other test words; it does not reflect a difference in the effect of test position on the referent and nonreferent test words. Although the referent does not slow as much from first to second test position (22 ms) as the nonreferent (41 ms), suggesting relative facilitation for the referent, the difference was not significant by post hoc tests.

## Experiments 3 and 4

In Experiments 1 and 2, the main result is a null result: moving from the test position before the pronoun to test positions after the pronoun did not produce any significant facilitation of the referent test word relative to the nonreferent test word. This lack of effect is consistent with the hypothesis that processing of the pronoun does not distinguish between the two characters; we would attribute this to the two characters being equally in the focus of attention. However, before accepting the null result, we tested

it further in Experiments 3 and 4.

In Experiment 3, all three of the test positions used in Experiments 1 and 2 were combined in one experiment. The materials and procedure were the same as in Experiments 1 and 2. As in Experiments 1 and 2, the results of Experiment 3 showed no significant facilitation of the referent relative to the nonreferent as test position moved from before the pronoun to the test positions after the pronoun. With a total of 117 subjects, this finding seems conclusive.

The finding is inconsistent with the results of past experiments (Chang, 1980; Corbett & Chang, 1983; Gernsbacher, 1989) in which referent test words were significantly facilitated over nonreferent test words. One possible reason for the difference in results is the reason suggested in the introduction: different kinds of processing may have occurred in our experiments than in the previous experiments. The faster reading times and response times we used may have led to exclusively automatic processing of pronouns, and the slower reading times and response times in the earlier experiments may have led to more strategic processing. The only directly comparable previous research that might have used an equivalently fast presentation rate (MacDonald & MacWhinney, 1990, in which the auditory presentation rate was not specified) did not obtain consistent results across two experiments. In one of their experiments, response times to a referent probe were faster than response times to a nonreferent probe when they were tested immediately after the pronoun, but in a second experiment response times to the two probes did not differ when immediately tested. Also, differences between referent and nonreferent response times at later test points were due in one experiment to a relative slow down of the nonreferent response times from immediate testing to later testing; in the other experiment, they were due to a speed up of the referent. A further difference between past experiments and ours is that we used comprehension questions that tested a variety of kinds of information from the texts. In earlier experiments, the comprehension questions usually required identification of the intended referent for the pronoun by asking subjects to verify which character performed the action of the final clause. Like the slow reading times, these questions may have encouraged strategic kinds of processing during reading.

However, a difference in kind of processing is not the only possible reason for the discrepancy between the results of Experiments 1 through 3 and earlier results. Another possibility might arise from the fact that the pronoun in the final clause in our experiments was always intended to refer to the character that was the subject of the first clause. In other studies, the pronoun sometimes referred to the subject and sometimes to the object. So in Experiment 4, we changed half of our materials to make the object of the first

clause the intended referent. It is also possible that there is some other unidentified difference between our materials and those used previously that is relevant to pronoun comprehension. To check this possibility, we included in Experiment 4 a small set of materials from experiments by Gernsbacher (1989). However, just as in the preceding experiments, the data show no significant differences between referent and nonreferent test word responses as a function of test position.

## Experiments 5 and 6

The conclusion from Experiments 1 through 4 is clear: for the sentences used in the experiments, referents and nonreferents are not differentially affected by processing of the pronoun. This conclusion holds over 157 subjects, over referents expressed as subjects and referents expressed as objects, over our materials as well as a subset of Gernsbacher's (1989) materials, and over cumulative and non-cumulative procedures for presenting texts.

Our interpretation of this result is that subjects were engaging in sentence processing that does not require the referent of the pronoun to be uniquely identified. For the sentences of the experiments, both characters are about equally in the discourse focus of attention, and information in the pronoun's clause is attached to the focus, not to either of the characters individually. Therefore, neither character gains in accessibility relative to the other. From this interpretation, we can make two testable predictions. First, if we can change subjects' processing to the appropriate strategies, the intended referent should be uniquely identified and we should see a relative advantage of referent over nonreferent test words. This was the aim of Experiments 5, 6, and 7. Second, we should be able to contrast the pronominal anaphors that are not uniquely identified with other kinds of anaphors for which the referent is identified. We do this in Experiments 8 and 9.

To encourage subjects to adopt a strategy of identifying the referents of the pronouns during reading, we needed to give them motivation to do the appropriate processing; we needed to make it relatively easy for them to do it; and we needed to give them time to do it. To provide motivation, each text was followed by a comprehension question for which the answer required that the actor of an action in the final sentence be identified. For the experimental sentences, this always required that the referent of the pronoun in the final clause be identified. To make it easy to do the appropriate processing, we used texts of only one sentence (for the experimental texts, this was the final sentence), so that subjects would know exactly what information the comprehension question would ask about and when to expect the pronoun in the text. To give subjects time to compute the intended referents of the pronouns, we adopted the procedure

used by Gernsbacher (1989), in which the time available for processing each word was 450 ms, plus 16 2/3 ms multiplied by the number of letters in the word. With this procedure, Gernsbacher found a large relative advantage of referents over nonreferents at the end of the sentence, and we expected to replicate this effect.

In Experiment 5, the referent and nonreferent character names were tested either immediately before the pronoun or at the end of the sentence. As expected, we found a larger relative advantage for the referent test word over the nonreferent test word at the end of the sentence than before the pronoun, indicating that our efforts to change subjects' processing were successful. The advantage came from an increase in response times for the nonreferent test words, which is consistent with Gernsbacher's hypothesis that processing of the pronoun gives an advantage to the referent by suppressing the nonreferent. However, as discussed in the introduction, this hypothesis can be tested with a control word. If suppression affects only the nonreferents, then the nonreferents should increase in response time at the end of the sentence relative to the referent, but the control word should not. This prediction was tested in Experiment 6, and it was not supported; response times for the control word did increase at the end of the sentence.

In contrast to Experiments 1 through 4, the results of Experiment 5 showed a relative advantage for referents over nonreferents. We attribute this advantage to pronominal processing that occurred because subjects were encouraged by the experimental procedure to identify the pronouns' referents during reading. Our interpretation of our results is that, with the same set of materials, processing can be exclusively automatic, leaving the pronoun unresolved (as in Experiments 1 - 4), or it may also include slower, strategic processes that allow the unique identification of the pronoun's referent (Experiment 5).

The results of Experiment 6 suggest reformulation of the suppression hypothesis proposed by Gernsbacher (1989). While we replicated the result that nonreferent response times were slower after the pronoun, responses for control words were slowed at least as much. This could be because suppression affects all entities in the discourse model (other than the referent). Or it could be that all test words are slowed because of end-of-sentence processing, and the underlying mechanism for the referent-nonreferent difference is actually facilitation for the referent. Currently, this issue cannot be resolved, and further research is needed.

## Experiment 7

In Experiment 5, strategic processing was encouraged by providing motivation to identify pronominal referents, by providing a sufficiently slow rate of presentation for the text, and by making the

task relatively easy with only one pronoun to be identified in a one-sentence text. The result was that referents showed a relative advantage over nonreferents, in contrast to Experiments 1 through 4. It might be thought that the only one of the three factors that actually contributed to the difference in findings between the first four experiments and Experiment 5 was the speed of presentation. Automatic processes of identification for the pronominal referents in the experimental texts might require more time than was available at the 250 ms per word rate used in the first four experiments. According to this hypothesis, simply slowing the rate of presentation should lead to an advantage for referents over nonreferents.

In Experiment 7, we tested this hypothesis by replicating Experiment 2 with a slow rate of presentation. The materials were the same multi-sentence texts as were used in Experiment 2, but the rate was slowed to 450 ms per word plus 16 2/3 ms multiplied by the number of letters in the word, the same rate used in Experiment 5.

The data showed clearly that, in this experiment, slowing the rate of presentation did not lead to an advantage for the referent over the nonreferent following reading of the pronoun. There was no advantage even though the rate was extremely slow, and comprehension questions asked for specific knowledge of the pronoun's referent.

Why do subjects appear to identify the pronominal referent in Experiment 5 but not in Experiment 7? The procedural differences in the two experiments are the number of sentences in the texts, one sentence in Experiment 5, three sentences in Experiment 7, and the inclusion of the control test words in Experiment 7. But these differences, especially the first, are critical. With only one sentence, a reader can easily anticipate exactly when the pronoun will occur, and exactly what the comprehension question must be. Also, in Experiment 5, all the test words were names, so that it would make sense for readers to keep track carefully of "who did what." In Experiment 7, it would theoretically be possible to anticipate exactly when the critical pronoun would occur and exactly what the comprehension question would be, but to do this, readers would have to count the sentences as they read to know which was the third, and then anticipate the comprehension question. In short, Experiment 7 reduces the ability of subjects to engage in strategic processing compared to Experiment 5.

## Experiments 8 and 9

In Experiment 5, we were able to show that subjects could, under the appropriate conditions, identify the intended referents for the pronouns in the experimental sentences. However, we are still left

with a null result for the procedure used in Experiments 1 through 4, for which we claim that fast, automatic processing leaves the pronoun unresolved. In Experiments 8 and 9, we show that this procedure does allow identification of the referent for another type of anaphor. The fact that at least one kind of referent is identified shows that the 250 ms per word reading rate used in our experiments is not so fast that it prevents the comprehension of all kinds of implicit information.

The anaphors we used were the nominals from studies by Dell et al. (1983). An example is shown in Table 6. In the first version of the fourth sentence, the nominal *the criminal* is intended to refer to the burglar mentioned in the first sentence. In the other version, the subject noun phrase is not intended to refer to the burglar. Dell et al., using the same procedure as in Experiments 1 through 4, showed that when the referent was presented as a test word after the anaphor, response time was facilitated relative to when it was presented after the control noun phrase. From this result (and appropriate control conditions), Dell et al. concluded that comprehension of the anaphor involved identification of its referent. Dell et al. also tested an associate of the referent (e.g. *garage* for the text in Table 6); this test word was a word that had occurred in the first sentence of the text with the referent. When this word was presented immediately after the anaphor, it also showed facilitated response time relative to the control condition, indicating that processing of the anaphor increased the accessibility not only of the referent but also concepts associated with the referent.

In Experiment 8, we mixed the texts of the pronominal anaphors from Experiments 1 through 7 with the texts of nominal anaphors used by Dell et al., and tested the referent of the nominal (e.g. *burglar*). Experiment 9 was similar, except that we tested for both the referent and the associated concept from the first sentence (e.g. *garage*). The prediction was that results for both sets of texts would replicate what had been found previously: relative facilitation would be observed with the nominals (and the concepts associated with them) but not with the pronouns.

For the pronoun materials, once again referent and nonreferent response times were not differentially affected by test position. In contrast, the nominal anaphors showed significant facilitation for their referents and for concepts associated with their referents. In general, the pattern of data for the nominal anaphors closely replicates the pattern obtained by Dell et al.

The results of these experiments were exactly as predicted: at a relatively fast presentation rate, in the absence of comprehension questions designed to motivate identification of anaphoric referents during reading, recognition responses for referents were facilitated for the nominal anaphors but not for the

pronominal anaphors in the experimental materials. Our interpretation of these results is that the referent of a nominal anaphor was uniquely identified during reading, but that the referent of a pronoun was not. We interpret the results for the nominal anaphors as showing referent identification in the light of several converging pieces of data. First, the relative facilitation for the referent test word (*burglar*) might be thought due solely to the semantic relation with the anaphor (*criminal*), but this can not be the case because the associated test word (*garage*) also shows facilitation. Second, the relative inhibition in the control condition might be thought due to the introduction of a new concept (*cat*), but such inhibition would also be expected to appear on responses to test words other than the referent and it did not (Dell et al., 1983).

There are several reasons why the referent of a nominal might have been identified under the same conditions in which the referent of a pronoun was not. One possibility is that the nominal was a word semantically related to its referent, and the pronoun was not (except with respect to gender). It has been suggested that semantic relatedness is a general aid to inference processes because semantic information is easily and quickly available during processing (McKoon & Ratcliff, 1989a, 1989b). Another possibility, suggested by Gernsbacher (1989), is that the nominal is more specific than the pronoun. The nominal might contain such specific information that, in the relevant discourse, no discourse entity other than the intended referent matches the nominal to any degree at all. For example, the nominal *criminal* may contain information specific enough that only *burglar* and no other entities in the discourse (such as *banker*) match the nominal to any degree. Finally, it could be that the nominal provides a second repetition of its referent entity in a way that a pronoun does not (i.e., the nominal may add information about the entity to its discourse representation). Obviously, more research will be needed to distinguish among these possibilities. However, the contrast between processing of the nominal and pronominal anaphors does make clear one point: it makes little sense to ask whether a reader "understands" a discourse overall and in general; under the same contextual conditions, a reader may identify a unique referent for one kind of anaphor but not for another. Empirical investigations of discourse comprehension can only be made up of tests of the many individual processes that may or may not, depending on experimental and contextual conditions, constitute comprehension.

## Discussion

Our conclusion that people do not always identify a unique referent for a pronoun, while consistent with current discourse models, stands in contrast with previous work. Hence, we should consider the reasons we have come to a different conclusion than have previous researchers. In empirical terms, our conclusion

was different because our procedures for testing pronoun resolution were different. More importantly, our procedures were motivated by a different theoretical view than has previously guided psycholinguistic research on pronoun resolution. Representing a text as a discourse model entails consideration of the relative accessibilities of the entities in the model. In this context, a pronoun is viewed as a cue to one or more of the entities. This "pronoun as cue" notion naturally suggests the parallel access matching process assumed by current memory models. These models distinguish automatic processes from strategic processes, and our experiments were designed to examine the identification of referents as an automatic process.

To move readers away from special strategies brought about by task demands that might have occurred in previous studies, we introduced three major methodological modifications. First, our texts were presented at a rate of 250 ms/word, compared to an average of about 500 ms/word in some previous work (e.g., Gernsbacher, 1989). Second, our texts contained three sentences (compared to the single sentence used by other researchers) and multiple test points throughout the texts. Third, comprehension questions presented after the texts tested a variety of kinds of information in our experiments, whereas previous experiments often asked specifically for information about the intended referents of pronouns. These three changes were introduced to discourage subjects from engaging in strategic processes during reading to identify the pronouns. Avoiding strategic processing is important because of the nature of the question we are studying. We are not asking whether people *can* uniquely identify referents for pronouns, but whether they *automatically* do so during comprehension and whether they always do so. It is clear that readers are capable of uniquely identifying pronominal referents; what is less clear is whether it is always a part of the processes of comprehension.

In our efforts to eliminate strategic processing of pronouns, we might have used reading times so fast that readers engaged in no processing at all. However, the reading rates that we used were appropriate for our subject population. As Experiments 8 and 9 demonstrate, the same subjects reading at the same speed did appear to resolve other types of anaphors. Furthermore, a slower reading rate, by itself, was not sufficient to guarantee resolution of the pronominal anaphors in our experiments. We found facilitation of pronominal referents over nonreferents only when the slow rate was combined with motivation to identity uniquely the referents and with procedures that made the identification task relatively easy.

Throughout the experiments described here, the distinction between automatic and strategic processes was used to guide choices of experimental variables. The application of the automatic/strategic distinction to reading processes is, of course, not straightforward. However, in some sense, the distinction

must apply; in reading, as in other cognitive tasks, there are processes that are slow and invoked to meet specific contextual demands, and there are processes that are faster and less constrained by a particular context. And the distinction can usefully be applied even though there are many open questions, such as whether the distinction represents a dichotomy or a continuum, and how the particular variables and results found for automatic processes in other domains can be applied to reading.

The usefullness of the distinction is demonstrated by the outcomes of the experiments. The distinction suggests experiments designed to move processing away from strategies adopted for a particular experimental task. Such strategies are generally assumed to be slower and more influenced by specific task demands than automatic processes and so, to eliminate them, reading and response rates were speeded and task demands specific to anaphoric identification were eliminated. Clearly, if there is a distinction (or a continuum) between automatic and strategic processes in reading, these procedural changes should represent a move towards the automatic. The fact that these procedural changes brought about substantial changes in the results of the experiments gives support to the utility of the automatic/strategic distinction in investigations of reading. The support for the automatic/strategic distinction is particularly impressive because it is only this notion, and not other current views, that would have guided us to address these questions in these ways. Previous views would have labelled anaphor resolution a necessary part of reading, and would not have suggested that anaphor resolution would depend on manipulations of task demands and rate of processing, except as part of a general failure in processing. Thus, the automatic/strategic distinction led to experiments that would otherwise not have been conducted, yet demonstrate important and unexpected boundary conditions on a fundamental aspect of reading.

By adopting the procedural manipulations suggested by an automatic/strategic distinction, we showed that the advantage in testing for the referent of a pronoun over a nonreferent could be eliminated. We interpret this result as indicating that the referent did not enjoy a processing advantage during reading over the nonreferent, and as providing support for the discourse model framework proposed in our research. According to this framework, the referent has no advantage because it was not uniquely identified as the referent of the pronoun.

An alternative interpretation is that the referent of the pronoun was, in fact, identified, but that this identification process did not lead to an advantage on the recognition test. One obvious possible reason for this would be that responses on the recognition test were at ceiling, but responses in Experiment 7 were relatively slow, yet still showed no facilitation for the referent. Other reasons that recognition might fail to

show the consequences of identification would be less plausible. In identification, the comprehension system must by some mechanism choose between two possible referents (e.g., *John* and *Mary*) on the basis of sex. Then, after making a choice, the system must either create a new token of the referent to which to attach the information given with the pronoun or attach the new information to the referent directly. Either way, new information about the referent would be encoded in memory. Thus, resolving the pronoun would entail both choosing the referent and encoding additional information about it, and this processing would have to be assumed to leave no consequences detectable in the recognition test.

Furthermore, assuming that identification leaves no traces detectable by recognition probes runs counter to all current accounts of on-line recognition testing (Chang, 1980; Corbett & Chang, 1983; Gernsbacher, 1989; MacDonald & MacWhinney, 1990; McKoon & Ratcliff, 1986; 1990; van Dijk & Kintsch, 1983). A variety of similar on-line processes are frequently observed on recognition tests. Experiments 8 and 9 present one example, where the effects of processing a noun anaphor are observed. Other examples include the processing of explicitly mentioned entities (Caplan, 1972; Jarvella, 1971), the processing of pronouns in object case ("him," "her"; Cloitre & Bever, 1989), the processing of empty syntactic traces (Bever & McElree, 1988), the processing of pronouns that refer to entities introduced in previous sentences (McKoon, Ward, Ratcliff, & Sproat, submitted), and the processing of verbs that take implicit instruments (McKoon & Ratcliff, 1981). Collectively, these examples overlap with the experiments reported here in many ways. The distance, in terms of number of words, between pronoun and antecedent is about the same in the current experiments as in the experiments of McKoon and Ratcliff (1980; two sentence texts), McKoon et al. (submitted), and Bever and McElree (1988). The type of pronoun (subject of its clause) is the same as in McKoon et al. (submitted). The use of the referent as test word is the same as in McKoon and Ratcliff (1980; 1981) and McKoon et al. (submitted). In all of these cases, processing facilitates recognition responses for the referenced entity. The only apparent difference in the experiments reported here is the presence of two possible referents for the pronoun.

We believe that the more plausible interpretation of the data is that the referent of the pronoun is not uniquely identified; instead, information given with the pronoun is attached to the current focus of attention, which includes both potential referents. One way that this could come about is suggested by current discourse models.

Discourse models have been proposed to describe the information that is used to establish co-reference among discourse entities. For a discourse model, the important variables that distinguish entities

are their relative accessibilities and their semantic (and possibly pragmatic) content. Variables such as recency of mention in the text and syntactic category are relevant only in their indirect effects on accessibility. More directly relevant are variables such as the relation between an entity and the discourse topic (Kintsch, 1974; McKoon et al., 1990), and variables that affect the semantic overlap among the entities. For example, reference processes can be affected by the degree of semantic association between an anaphor and its possible referents (Corbett, 1984).

A model of discourse processing in which pronouns are matched against all entities in memory suggests that there may be some contexts in which no single discourse entity matches sufficiently better than all others to be selected as the referent. In the experiments presented here, it appears that we have found one set of contextual factors in which that happens. However, we would be ill-advised to conclude that this situation is the general one, or even a common one. We have only studied texts with two relatively indistinguishable characters, one of whom is referred to by a pronoun. In fact, much of the research on pronoun comprehension consists of studies using materials that fit the same general description (Chang, 1980; Corbett & Chang, 1983; Ehrlich, 1980; Garnham & Oakhill, 1985; Gernsbacher, 1989; MacDonald & MacWhinney, 1990). But this is far from the situation in which we would expect pronouns to occur most often in natural discourses. Normally, when a pronoun is used, one discourse entity is already in the focus of attention (Brennan, 1989; Chafe, 1974; Fletcher, 1984). It seems that we have been studying pronouns outside their natural habitat.

Moreover, it may be that pronouns have been studied for the wrong reasons. In past studies, the problem has been to find out how the processing system uses a pronoun to find its referent. Phrasing the question this way puts the burden on processes driven by the pronoun. But the appropriate question may be to ask not what the pronoun does for the discourse, but what the discourse does for the pronoun. When the discourse has only one entity in the focus of attention at the time the pronoun is encountered, then it may be that essentially no processing is required for the pronoun. It may be that information predicated of the pronoun is attached to the focussed entity via an attachment process that is simple, automatic, and demanding of little processing capacity. If this is the case, then pronouns are interesting not because of the effort they require but precisely because of the effort they do not require.

Our suggestion is that pronouns are most frequently dealt with by an automatic process of attaching their propositions to the current discourse focus and the propositions relevant to it. It follows from this idea that the referent of a pronoun will be completely and correctly identified only if the discourse focus contains

the uniquely correct referent. If the focus contains more than one possible referent, as in our experiments, then the propositions of the pronoun are attached equally to all the focused entities. In effect, the automatic processes of comprehension treat the new information simply as predicated of the entity or entities in focus. This processing may not always result in the "correct" representation of a text in some ultimate sense for some particular set of experimental materials; instead the processing system is designed to operate under stringent time constraints to provide a useful understanding of natural discourse. Of course, if a comprehender has special motivation and enough time to resolve a pronoun reference more completely, he or she can engage in further, strategic, processing to do so.

Viewing pronouns as cues to discourse entities is consistent with three phenomena previously pointed out by other researchers: pronouns that refer via demonstration, "unheralded pronouns" (see Gerrig, 1986), and "conceptual anaphors" (see Gernsbacher, 1986). First, if a discourse is about some unique but linguistically unspecified referent, then the lack of linguistic specification does not necessarily impede comprehension. This has been documented by Clark, Schreuder, and Buttrick (1983), who note that linguistically underdetermined noun phrases can be used to refer to unstated entities that are nevertheless in common ground. For example, the assertion, "They publish gossip," uttered while pointing to a newspaper, refers successfully to the newspaper's publishers. Theories of pronoun resolution that conceive of pronouns as triggering a search for a linguistic referent cannot explain this example. In contrast, such examples fit naturally into a theory such as ours that views a pronoun as a cue relevant to some entity in the comprehender's discourse model. Reference by demonstration may not be understood by entirely automatic processes, yet whatever the processing, the result is resolution of an anaphor as referring to a focussed entity.

Unheralded pronouns (Gerrig, 1986) are also consistent with the pronoun-as-cue framework. An unheralded pronoun refers to an entity not previously referred to either linguistically or deictically. Consider the following conversation between two popular music buffs:

Penny: Do you have a CD of "Abbey Road?"
Cindy: Oh, sure. I have CD's of all their stuff.

For these speakers (and perhaps for some readers of this report), the pronoun *their* refers successfully to the Beatles. The pronoun-as-cue framework can account for this example by assuming that the album title brings the concept of the Beatles into the comprehender's discourse model, making it sufficiently accessible for the pronoun to be uttered felicitously.

The third phenomenon that can be understood from the pronoun-as-cue framework is what Gernsbacher (1986) refers to as "conceptual anaphora." Normally, pronouns in English agree in number with their referents. However, Gernsbacher notes exceptions such as the following:

I need a plate. Where do you keep them?

For examples such as this, in which the speaker is referring to an unspecified member of a set of items that all will serve equally well, the plural pronoun is rated as being more natural and is comprehended more quickly than the singular pronoun. Again, a traditional view of pronoun resolution would have difficulty explaining this phenomenon. However, the pronoun-as-cue framework simply assumes that the speaker's use of the word *plate* focuses the comprehender's attention on all of his or her plates. In this context, it is natural to refer to the entire set of plates via a pronoun.

As illustrated by these examples, the pronoun-as-cue framework encourages us to look at the larger discourse context to understood how pronouns are used felicitously. Pronouns are viewed as doing little more than signaling the comprehender that the speaker (or author) is referring to whatever entity is in the current focus of attention, within the constraints imposed by syntax. In this view, the interesting questions for research concern how various discourse elements are deployed to help the speaker (or author) and comprehender share the same focus of attention. To answer these questions, it will be necessary to look beyond the literal text of a discourse.

## References

Bever, T. & McElree, B. (1988). Empty categories access their antecedents during comprehension. *Linguistic Inquiry, 19*, 35-43.

Brennan, S.E. (1989). Centering attention in discourse. Unpublished manuscript, Stanford University.

Caplan, D. (1972). Clause boundaries and recognition latencies for words in sentences. *Perception and Psychophysics, 12*, 73-76.

Chafe, W.L. (1974). Language and consciousness. *Language, 50*, 111-133.

Chafe, W.L. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C.N. Li (Ed.), *Subject and topic* (pp. 25-55). New York: Academic Press.

Chang, F.R. (1980). Active memory processes in visual sentence comprehension: Clause effects and pronominal reference. *Memory and Cognition, 8*, 58-64.

Clark, H.H., & Marshall, C.R. (1981). Definite reference and mutual knowledge. In A.K. Joshi, B.L. Webber, and I.A. Sag (Eds.), *Elements of discourse understanding*. Cambridge: Cambridge University Press, 10-63.

Clark, H.H., Schreuder, R., & Buttrick, S. (1983). Common ground and the understanding of demonstrative reference. *Journal of Verbal Learning and Verbal Behavior, 22*, 245-258.

Clark, H.H., & Sengul, C.J. (1979). In search of referents for nouns and pronouns. *Memory and Cognition, 7*, 35-41.

Clifton, C., Jr., & Ferreira, F. (1987). Discourse structure and anaphora: Some experimental results. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 635-654). Hillsdale, NJ: Erlbaum.

Corbett, A.T. (1984). Prenominal adjectives and the disambiguation of anaphoric nouns. *Journal of Verbal Learning and Verbal Behavior, 23* 683-695.

Corbett, A.T., & Chang, F.R. (1983). Pronoun disambiguation: Accessing potential antecedents. *Memory and Cognition, 11*, 283-294.

Dell, G.S., McKoon, G., & Ratcliff, R. (1983). The activation of antecedent information during the processing of anaphoric reference in reading. *Journal of Verbal Learning and Verbal Behavior, 22*,

121-132.

van Dijk, T.A., & Kintsch, W. (1983). *Strategies of discourse comprehension.* New York: Academic Press.

Ehrlich, K. (1980). Comprehension of pronouns. *Quarterly Journal of Experimental Psychology, 32,* 247-255.

Ehrlich, K. (1983). Eye movements in pronoun assignment: A study of sentence integration. In K. Rayner, (Ed.), *Eye movements in reading: Perceptual and language processes* (pp. 253-272). New York: Academic Press.

Ehrlich, S.F., & Rayner, K. (1983). Pronoun assignment and semantic integration during reading: Eye movements and immediacy of processing. *Journal of Verbal Learning and Verbal Behavior, 22,* 75-87.

Fletcher, C. (1984). Markedness and topic continuity in discourse processing. *Journal of Verbal Learning and Verbal Behavior, 23,* 487-493.

Fodor, J.D. (1989). Empty categories in sentence processing. *Language and Cognitive Processes, 4,* 155-209.

Garnham, A., & Oakhill, J.V. (1985). On-line resolution of anaphoric pronouns: Effects of inference making and verb semantics. *British Journal of Psychology, 76,* 385-393.

Garvey, C., Caramazza, A., & Yates, J. (1976). Factors influencing assignment of pronoun antecedents. *Cognition, 3,* 227-243.

Gernsbacher, M. (1986). The comprehension of conceptual anaphora in discourse. *Proceedings of the Eighth Annual Conference of the Cognitive Science Society,* 110-125.

Gernsbacher, M.A. (1989). Mechanisms that improve referential access. *Cognition, 32,* 99-156.

Gernsbacher, M. & Hargreaves, D. (1988). Accessing sentence participants: The advantage of first mention. *Journal of Memory and Language, 27,* 699-717.

Gernsbacher, M., Hargreaves, D. & Beeman, M. (1989). Building and accessing clausal representations: The advantage of first mention versus the advantage of clause recency. *Journal of memory and Language, 28,* 735-755.

Gerrig, R.J. (1986). Process models and pragmatics. In N.E. Sharkey (Ed.), *Advances in cognitive science* (pp. 23-42). Chichester, England: Ellis Horwood.

Gillund, G. & Shiffrin, R.M. (1984). A retrieval model for both recognition and recall. *Psychological Review, 91*, 1-67.

Givon, T. (1976). Topic, pronoun, and grammatical agreement. In C.N. Li (Ed.), *Subject and topic* (pp. 149-188). New York: Academic Press.

Grosz, B. (1977). The representation and use of focus in a system for understanding dialogs. In *Proceedings of the Fifth International Joint Conference on Artificial Intelligence*, Los Altos: William Kaufmann, Inc.

Grosz, B. (1981). Focusing and description in natural language dialogues. In A.K. Joshi, B.L. Webber, and I.A. Sag (Eds.), *Elements of discourse understanding* (pp. 84-105). Cambridge, MA: Cambridge University Press.

Grosz, B.J., Joshi, A.K., & Weinstein, S. (1983). Providing a unified account of definite noun phrases in discourse. *Proceedings of the 21st Annual Meeting of the Association of Computational Linguistics*, 44-50.

Grosz, B. & Sidner, C. (1986). Attention, intentions and the structure of discourse. *Computational Linguistics, 12*, 175-204.

Haviland, S.E., & Clark, H.H. (1974). What's new? Acquiring information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior, 13*, 512-521.

Hawkins, J.A. (1977). The pragmatics of definiteness I and II. *Linguistiche Berichte, 47*, 1-27.

Hintzman, D. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review, 95*, 528-551.

Hobbs, J. (1978). Resolving pronoun references. *Lingua, 44*, 311-338.

Hudson, S.B., Tanenhaus, M.K., & Dell, G.S. (1987). The effect of the discourse center on the local coherence of a discourse. *Proceedings of the Eighth Annual Conference of the Cognitive Science Society*, 96-101.

Jarvella, R.J. (1971). Syntactic processing of connected speech. *Journal of Verbal Learning and Verbal Behavior, 10*, 409-416.

Just, M.A., & Carpenter, P.A. (1980). A theory of reading: From eye fixations to comprehension.

*Psychological Review*, *87*, 329-354.

Just, M.A., & Carpenter, P.A. (1987). *The psychology of reading and language comprehension.* Boston, MA: Allyn and Bacon.

Karttunen, Lauri. 1976. Discourse referents. Syntax and semantics VII: Notes from the linguistic underground, ed. by James McCawley, 363-86. New York: Academic Press.

Kintsch, W. (1974). *The representation of meaning in memory.* Hillsdale, NJ: Erlbaum.

Kintsch, W. & van Dijk, T.A. (1978). Toward a model of text comprehension and production. *Psychological Review*, *85*, 363-394.

Kintsch, W., & Keenan, J.M. (1973). Reading rate and retention as a function of the number of propositions in the base structure of sentences. *Cognitive Psychology*, *5*, 257-274.

Levi, Judith N. 1978. The syntax and semantics of complex nominals. New York: Academic Press.

Lieber, Rochelle. 1984. Grammatical rules and sublexical elements. Papers from the parasession on lexical semantics, Chicago Linguistic Society 20.187-199.

MacDonald, M.C., & MacWhinney, B. (1990). Measuring Inhibition and Facilitation from Pronouns. *Journal of Memory and Language*, *29*, 469-492.

Matthews, A. & Chodorow, M. (1988). Pronoun resolution in two-clause sentences: Effects of ambiguity, antecedent location and depth of embedding. *Journal of Memory and Language*, *27*, 245-260.

McKoon, G. (1977). Organization of information in text memory. *Journal of Verbal Learning and Verbal Behavior*, *16*, 247-260.

McKoon, G., & Ratcliff, R. (1980). The comprehension processes and memory structures involved in anaphoric reference. *Journal of Verbal Learning and Verbal Behavior*, *19*, 668-682.

McKoon, G., & Ratcliff, R. (1980a). Priming in item recognition: The organization of propositions in memory for text. *Journal of Verbal Learning and Verbal Behavior*, *19*, 369-386.

McKoon, G., & Ratcliff, R. (1981). The comprehension processes and memory structures involved in instrumental inference. *Journal of Verbal Learning and Verbal Behavior*, *20*, 671-682.

McKoon, G., & Ratcliff, R. (1984). Priming and on-line text comprehension. In D.E. Kieras & M.A. Just (Eds.), *New methods in reading comprehension research* (pp. 119-128). Hillsdale, NJ: Erlbaum.

McKoon, G., & Ratcliff, R. (1986). Inferences about predictable events. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*, 82-91.

McKoon, G. & Ratcliff, R. (1989a). Inferences about contextually-defined categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 1134-1146.

McKoon, G., & Ratcliff, R. (1989b). Semantic association and elaborative inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 326-338.

McKoon, G. & Ratcliff, R. (1990). Textual inferences: Models and measures. In Balota, D.A., Flores d'Arcais, G.B., & Rayner, K. (Eds.), *Comprehension processes in reading* (pp. 403-421). Hillsdale, N.J.: Erlbaum.

McKoon, G., & Ratcliff, R. Inference during reading. Submitted.

Morgan, Jerry. (1978). Toward a rational model of discourse comprehension. Proceedings of TINLAP-2, ed. by David Waltz. 109-14. New York: Association for Computational Machinery

Murdock, B.B. (1974). *Human memory: Theory and data.* Potomac, MD: Erlbaum.

Murdock, B.B. (1982). A theory for the storage and retrieval of item and associative information. *Psychological Review, 89*, 609-626.

Neely, J.H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited capacity intention. *Journal of Experimental Psychology: General, 106*, 226-254.

Nicol, J., & Swinney, D. (1989). The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research, 18*, 5-20.

Posner, M.I., & Snyder, C.R. (1975). Attention and cognitive control. In R.L. Solso (Ed.), *Information processing and cognition* (pp. 55-85). Hillsdale, NJ: Erlbaum.

Postal, P. (1969). Anaphoric islands. In R. Binnick, et al. (eds.) *Papers from the fifth regional meeting of the Chicago Linguistic Society*. 205-255.

Prince, E. (1981). Toward a taxonomy of given-new information. In P. Cole (Ed.), *Radical pragmatics*, 223-255. New York: Academic Press.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*, 59-108.

Ratcliff, R. & McKoon, G. (1978). Priming in item recognition: Evidence for the propositional structure of

sentences. *Journal of Verbal Learning and Verbal Behavior, 17*, 403-417.

Ratcliff, R., & McKoon, G. (1981). Automatic and strategic priming in recognition. *Journal of Verbal Learning and Verbal Behavior, 20*, 204-215.

Ratcliff, R., & McKoon, G. (1988b). A retrieval theory of priming in memory. *Psychological Review, 95*, 385-408.

Rayner, K. (1978). Eye movements in reading and information processing. *Psychological Bulletin, 85*, 618-660.

Reinhart, T. (1982). *Pragmatics and linguistics: An analysis of sentence topics*. Bloomington, IN: Indiana University Linguistics Club.

Rothkopf, E., Koether, M., & Billington, M. (1988). *Why are certain sentence constructions mnemonically robust for modifiers?* Technical Memorandum, AT&T Bell Laboratories.

Sidner, C. (1983a). Focusing and discourse. *Discourse Processes, 6*, 107-130.

Sidner, C. (1983b). Focusing in the comprehension of definite anaphora. In M. Brady & R. Berwick (Eds.), *Computational models of discourse* (pp. 267-330). Cambridge, MA: MIT Press.

Simpson, Jane. (1983). Aspects of Warlpiri morphology and syntax. Cambridge, MA: MIT dissertation.

Tanenhaus, M.K., Carlson, G.N., & Trueswell, J.C. (1989). The role of thematic structures in interpretation and parsing. *Language and Cognitive Processes, 4*, 211-234.

Ward, Gregory L. (1985). The semantics and pragmatics of preposing. Philadelphia: University of Pennsylvania dissertation. *Reprinted in 1988, New York: Garland.*

Ward, G., Sproat, R., & McKoon, G. (1991). A pragmatic analysis of so-called anaphoric islands. *Language*,

Webber, B. (1979). Elements of discourse understanding. New York: Garland Press.

Webber, B. (1983). So what can we talk about now? In M. Brady and R. Berwick, (Eds.), *Computational models of discourse* (pp. 331-371). Cambridge, MA: MIT Press.

Wilson, D., & Sperber, D. (1979). Ordered entailments: An alternative to presuppositional theories. *Syntax and Semantics, 11*, 299-323.

Yule, G. (1982). Interpreting anaphora without identifying reference. *Journal of Semantics, 1*, 315-322.

## Articles:

Ward, G., Sproat, R., & McKoon, G. (1990). A pragmatic analysis of so-called anaphoric islands. *Language*, In press.

Ratcliff, R. & McKoon, G. (1991). Using ROC data and priming results to test global memory models. In S. Lewandowsky and W.E. Hockley (eds.), Human Memory, Theory and Data: Essays in Honor of B.B. Murdock, Jr. In press.

McKoon, G. & Ratcliff, R. (1990). Inference during reading. In press, *Psychological Review*.

Greene, S., McKoon, G., & Ratcliff, R. Pronoun resolution and discourse models. Submitted.

McKoon, G., Ward, G., Ratcliff, R., & Sproat, R. Morphosyntactic and pragmatic factors affecting the accessibility of discourse entities. Submitted.

## Papers Presented:

McMaster University, March, 1990.

CUNY Conference on Sentence Processing, March 1990.

Conference honoring B. Murdock, June, 1990.

Psychonomic Society Meeting, Nov, 1990.

Invited Address, CUNY Conference on Sentence Processing, May 1991.

## Table 1

**Examples of Texts Used in Experiment 1**

Referent noun: deer

High topicality, compound

Sam likes the outdoor life. Having grown up in rural Kentucky, he knows a lot about nature and is an expert at fishing and shooting.He goes on hunting trips as often as he can. He used to hunt just small game, like rabbit and quail. However, lately he's taken up deer hunting. He thinks that they are really exciting to track.

Low topicality, compound

Sam has many interests in the outdoors. He's an avid skier, and each winter he takes about a month off from work to ski in Colorado. In the summertime, he visits his parents in Montana where he has a chance to do some mountainclimbing. Lately, he's taken up deer hunting. He thinks that they are really exciting to track.

High topicality, verbal complement

Sam likes the outdoor life. Having grown up in rural Kentucky, he knows a lot about nature and is an expert at fishing and shooting. He goes on hunting trips as often as he can.He used to hunt just small game, like rabbit and quail.However, lately he's taken up hunting deer. He thinks that they are really exciting to track.

Low topicality, verbal complement

Sam has many interests in the outdoors. He's an avid skier, and each winter he takes about a month off from work to ski in Colorado. In the summertime, he visits his parents in Montana where he has a chance to do some mountain climbing. Lately, he's taken up hunting deer. He thinks that they are really exciting to track.

## Table 2

### Data from Experiment 1

Response Times and Error Rates for Test Words

| | Text Version | | | |
|---|---|---|---|---|
| | Low topicality | | High topicality | |
| Syntactic structure | | | | |
| Compound | 907 ms | 5 % | 870 ms | 2 % |
| Verbal complement | 893 ms | 4 % | 886 ms | 4 % |
| Filler positive test words | 1242 ms | 21 % | | |
| Filler negative test words | 1181 ms | 15 % | | |

Reading Times for Final Sentences

| | Text Version | |
|---|---|---|
| | Low topicality | High topicality |
| Syntactic structure | | |
| Compound | 2117 ms | 1785 ms |
| Verbal complement | 1868 ms | 1738 ms |

## Table 3

### Data from Experiment 2

Response Times and Error Rates for Test Words

| | Final Sentence | | | |
|---|---|---|---|---|
| | Pronoun | | New noun | |
| Syntactic structure | | | | |
| Compound | 948 ms | 7 % | 1070 ms | 8 % |
| Verbal complement | 926 ms | 5 % | 1045 ms | 10 % |
| Filler positive test words | 1263 ms | 23 % | | |
| Filler negative test words | 1150 ms | 14 % | | |

Reading Times for Final Sentences

| | Final Sentence | |
|---|---|---|
| | Pronoun | New noun |
| Syntactic structure | | |
| Compound | 1961 ms | 2199 ms |
| Verbal complement | 2012 ms | 2254 ms |

## Table 4

An Example of the Experimental Texts

Mary and John were doing the dishes after dinner.
One of them was washing while the other dried.
Mary accidentally scratched John with a knife
and then$_1$ she dropped$_2$ it on the counter.$_3$

Test words

Referent: Mary
Nonreferent: John
Control: dishes

# Table 5

Results of Experiments 1 and 2
Response Times and Error Rates on Test Words

| **Experiment 1** | Test Position | | | |
|---|---|---|---|---|
| | 1 | | 2 | |
| Test word | | | | |
| Referent | 656 | 7% | 669 | 10% |
| Nonreferent | 633 | 4% | 624 | 3% |
| Control | 729 | 12% | 746 | 15% |
| Positive fillers | 779 | | 11% | |
| Negative fillers | 832 | | 13% | |

| **Experiment 2** | Test Position | | | |
|---|---|---|---|---|
| | 1 | | 3 | |
| Test Word | | | | |
| Referent | 675 | 7% | 697 | 7% |
| Nonreferent | 654 | 5% | 695 | 2% |
| Control | 705 | 11% | 784 | 20% |
| Positive fillers | 711 | | 26% | |
| Negative fillers | 799 | | 15% | |

| **Procedure Check Experiment** | Test Position | | | |
|---|---|---|---|---|
| | 1 | | 3 | |
| Test Word | | | | |
| Referent | 721 | 8% | 731 | 8% |
| Nonreferent | 712 | 8% | 718 | 4% |
| Control | 785 | 15% | 845 | 24% |
| Positive fillers | 820 | | 22% | |
| Negative fillers | 829 | | 12% | |

## Table 6

An Example of the Paragraphs Used
in Experiments 8 and 9

Sentence 1: A burglar surveyed the garage set back from the street.

Sentence 2: Several milk bottles were piled at the curb.

Sentence 3: The banker and her husband were on vacation.

Sentence 4: (Version 1, anaphor):
The criminal slipped$_2$ away from the streetlamp.$_3$

Sentence 4: (Version 2, no anaphor):
A cat slipped$_2$ away from the streetlamp.$_3$

Test words
Referent: burglar

Associate of referent: garage

ADDENDUM

Papers referencing AFOSR sponsored research by collaborators and McKoon:

Sproat, Richard. A pragmatic analysis of so-called anaphoric islands, (with Gregory Ward and Gail McKoon), University of Pennsylvania, November 15, 1990.

Ward, Gregory. No Word Is an Island: A Pragmatic Analysis of Outbound Anaphora: Invited speaker at the workshop on the Pragmatics of Language, sponsored by the University of Chicago. April, 1991.

Ward, Gregory. A Pragmatic Analysis of Outbound Anaphora and "Vice Versa", Conference on Grammatical Foundations of Prosody and Discourse, sponsored by NSF. University of California, Santa Cruz. July, 1991. [with R. Sproat]